

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: September 15, 2004, 21:25:56 ; Search time 6853 Seconds  
(without alignments)  
6830.655 Million cell updates/sec

Title: US-10-759-277-3  
Perfect score: 1080  
Sequence: 1 atgcgcgtggagctggagct.....gcaaaaaaatcaagtcttaa 1080

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rod.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.hugo.hum.\*
- 40: em.hugo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1080	100.0	1080	6	BD270235	Human gly
2	1080	100.0	1080	9	AF257466	Human gly
3	1076.8	99.7	1181	9	BC019315	Homo sapi
4	1076.8	99.7	1229	9	BC000008	Homo sapi
5	1076.8	99.7	1230	6	AR339544	Sequence
6	1075.2	99.6	1170	6	AX877224	Sequence
7	1075.2	99.6	1170	6	BD156549	Primer fo
8	1075.2	99.6	1170	9	AK001659	Homo sapi
9	883.2	81.8	1247	10	BC003307	BC003307 Mus muscu
10	883.2	81.8	1247	10	BC057977	BC057977 Mus muscu
11	881.6	81.6	1174	10	AB041263	AB041263 Mus muscu
12	732.4	67.8	3026	9	AK091184	Homo sapi
13	710.2	65.8	813	6	AX868777	Sequence
14	710.2	65.8	813	6	BD148839	Primer fo
15	483.8	44.8	514	6	AX873806	Sequence
16	483.8	44.8	514	6	BD153868	Primer fo
17	456.6	42.3	237317	2	AC097420	Rattus no
18	418	38.7	464	6	AX408867	Sequence
19	386.8	35.8	553	6	AX557002	Sequence
20	338.6	31.4	228917	2	AC106079	Rattus no
21	338.6	31.4	229321	2	AC094185	Rattus no
22	280.6	26.0	404	6	AX201885	Sequence
23	269.8	25.0	2044	9	AK126144	Homo sapi
24	269.8	25.0	201300	9	AL137073	Human DNA
25	241.6	22.4	227816	10	AL683884	Mouse DNA
26	241.4	22.4	244134	2	AC097073	Rattus no
27	221	20.5	1194	3	AF397531	Drosophil
28	210.4	19.5	1999	9	AK123739	Homo sapi
29	199.6	18.5	197522	2	AC137022	Rattus no
30	195	18.1	182381	2	AC122081	Rattus no
31	195	18.1	191841	2	AC112827	Rattus no
32	195	18.1	266461	2	AC120718	Rattus no
33	145.6	13.5	292100	1	SC0939121	Streptomy
34	144.4	13.4	11237	1	AE012850	Chlorobiu
35	135	12.5	302300	1	AP005034	Streptomy
36	126.4	11.7	298050	1	EX321861	Nitrosomo
37	125	11.6	349652	1	EX569690	Synechoco
38	119.8	11.1	15267	3	AB055099	Drosophil
39	119.8	11.1	22990	3	AB055101	Drosophil
40	119.8	11.1	23007	3	AB055100	Drosophil
41	119.8	11.1	23244	3	AB055098	Drosophil
42	119.8	11.1	61204	2	AC017132	Drosophil
43	119.8	11.1	161601	3	AC007594	Drosophil
44	119.8	11.1	225655	3	AE003695	Drosophil
45	111.8	10.4	86896	1	RCU57682	Rhodobacter

ALIGNMENTS

RESULT 1  
BD270235  
LOCUS BD270235 1080 bp DNA linear PAT 17-JUL-2003  
DEFINITION Human glycosylation enzymes.  
ACCESSION BD270235  
VERSION BD270235.1 GI:33080003  
KEYWORDS JP 2002537796-A/2.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1080)  
AUTHORS Coleman,T.A. and Betenbaugh,M.J.  
TITLE Human glycosylation enzymes  
JOURNAL Patent: JP 2002537796-A 2 12-NOV-2002;

COMMENT	HUMAN GENOME SCIENCES INC, JOHNS HOPKINS UNIVERSITY
OS	Homo sapiens (human)
PN	JP 2002537796-X/2
PD	12-NOV-2002
PF	01-MAR-2000 JP 2000602748
PR	02-MAR-1999 US 60/122409
PI	TIMOTHY A COLEMAN, MICHAEL J BETENBAUGH
PC	C12N15/09, A61K31/711, A61K38/43, A61K48/00, A61P1/02, A61P3/00, PC A61P7/00, A61P19/02, A61P17/02, A61P15/00, A61P17/02, PC A61P19/02, A61P19/10, A61P21/04, A61P25/00, A61P25/16, A61P25/24, A61P25/28, PC A61P29/00, A61P31/18, A61P35/00, A61P37/00, A61P43/00, C07K16/40, C12N1/15, PC C12N1/19, PC C12N1/21, C12N5/10, C12N9/00, G01N33/15, G01N33/50, G01N33/53, PC C12N15/00, PC C12N5/00, A61K37/48
CC	Human glycosylation enzymes
FH	Key Location/Qualifiers
FT	CDS (1)..(1077).
FEATURES	Location/Qualifiers
source	1..1080
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
ORIGIN	
Query Match	100.0%; Score 1080; DB 6; Length 1080;
Best Local Similarity	100.0%; Pred. No. 2.2e-256;
Matches 1080; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ATGCGCTGGAGCTGCTGCTCCGGCGCTGGTGGGGGCGGCAACACCCGTCCTC 60
Db	1 ATGCGCTGGAGCTGCTGCTCCGGCGCTGGTGGGGGCGGCAACACCCGTCCTC 60
QY	61 ATCATTTCCGAGATCGGCAGAACACACAGGGCGACCTGGAGCTAGCCAGGCGATGATC 120
Db	61 ATCATTTCCGAGATCGGCAGAACACACAGGGCGACCTGGAGCTAGCCAGGCGATGATC 120
QY	121 CGATGGCCAGGAGTGTGGGCTGATTGTGCGCAAGTTCAGAGAGTGAGTAGAATTC 180
Db	121 CGATGGCCAGGAGTGTGGGCTGATTGTGCGCAAGTTCAGAGAGTGAGTAGAATTC 180
QY	181 AAGTTTAATCGAAGACCTTGAGAGGCCATACACTCGAAGCTTCCTGGGGGAGAGC 240
Db	181 AAGTTTAATCGAAGACCTTGAGAGGCCATACACTCGAAGCTTCCTGGGGGAGAGC 240
QY	241 TACGGGAGCAAAACGACATCTGAGATTTCAGCCATGACCATGACAGGAGCTGCAGAGG 300
Db	241 TACGGGAGCAAAACGACATCTGAGATTTCAGCCATGACCATGACAGGAGCTGCAGAGG 300
QY	301 TACGGGAGGAGTTGGGATCTTCTTCTACTGCTCTGGCATGGATGAGATGGCAGTTGAA 360
Db	301 TACGGGAGGAGTTGGGATCTTCTTCTACTGCTCTGGCATGGATGAGATGGCAGTTGAA 360
QY	361 TTCTGCGATGAATCAATGTTCCATTTTCAAAGTTGGATCTGGAGACACATAATATTT 420
Db	361 TTCTGCGATGAATCAATGTTCCATTTTCAAAGTTGGATCTGGAGACACATAATATTT 420
QY	421 CCTTATCTGGAAGACAGCAAAAGAGTTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
Db	421 CCTTATCTGGAAGACAGCAAAAGAGTTCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
QY	481 TCATGGACACCATGAAGCAGTTTATCAGATCGTGAAGCCCTCAACCCCACTTCTGC 540
Db	481 TCATGGACACCATGAAGCAGTTTATCAGATCGTGAAGCCCTCAACCCCACTTCTGC 540
QY	541 TTCTTGCAGTGTACCGAGGCATACCCGCTCCAGCTGAGGAGCTCAACCTGCGGTCTATC 600
Db	541 TTCTTGCAGTGTACCGAGGCATACCCGCTCCAGCTGAGGAGCTCAACCTGCGGTCTATC 600
QY	601 TCGGAATATCAGAGCTCTTTCTGACATTCCTCATAGGATATCTGGGCAATGAACAGGC 660
Db	601 TCGGAATATCAGAGCTCTTTCTGACATTCCTCATAGGATATCTGGGCAATGAACAGGC 660
QY	661 ATAGCGATATCTGTGCGCGCAGTGGCTCTGGGGGCGCAAGGTGTTGGAACGTACATAACT 720
Db	661 ATAGCGATATCTGTGCGCGCAGTGGCTCTGGGGGCGCAAGGTGTTGGAACGTACATAACT 720
QY	721 TTGGACACAGACTGGAAGGGAGTGACCACTGGGCTCTGGAGCTTGAGAGCTGGAGCTGGCC 780
Db	721 TTGGACACAGACTGGAAGGGAGTGACCACTGGGCTCTGGAGCTTGAGAGCTGGAGCTGGCC 780
QY	781 GAGCTGGTGGCTGAGTGGCTCTGTGAGCGCTGCGGCTCCCAACCAAGCAGCTG 840
Db	781 GAGCTGGTGGCTGAGTGGCTCTGTGAGCGCTGCGGCTCCCAACCAAGCAGCTG 840
QY	841 CTGCGCTGTGAGATGGCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGA 900
Db	841 CTGCGCTGTGAGATGGCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGA 900
QY	901 ATTCCGGAGGACCACTTCAACATGACATGCTCAGCGTGAAGGTGGTGGAGCCCAA 960
Db	901 ATTCCGGAGGACCACTTCAACATGACATGCTCAGCGTGAAGGTGGTGGAGCCCAA 960
QY	961 GCCTATCTCTGGAAGACATCTTTAATCTAGTGGGCAAGAGTCTGTGCTCACTGTGAA 1020
Db	961 GCCTATCTCTGGAAGACATCTTTAATCTAGTGGGCAAGAGTCTGTGCTCACTGTGAA 1020
QY	1021 GAGGATGACACATCATGGAAGATTGGTAGAATATCATGGCAAAATAAAGTCTTAA 1080
Db	1021 GAGGATGACACATCATGGAAGATTGGTAGAATATCATGGCAAAATAAAGTCTTAA 1080
RESULT 2	
AF257466	1080 bp mRNA linear PRI 11-JUN-2000
LOCUS	Homo sapiens N-acetylneuraminic acid phosphate synthase mRNA, complete cds.
ACCESSION	AF257466
VERSION	AF257466.1 GI:8453155
KEYWORDS	
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Lawrence, S.M., Huddleston, K.A., Pitts, L.R., Nguyen, N., Lee, Y.C., Vann, W.F., Coleman, T.A. and Betenbaugh, M.J.
TITLE	Cloning and expression of the human N-acetylneuraminic acid phosphate synthase gene with 2-keto-3-deoxy-D-glycero-D-galacto-nononic acid biosynthetic ability
JOURNAL	J. Biol. Chem. 275 (23), 17869-17877 (2000)
MEDLINE	20298869
PUBMED	10749855
REFERENCE	2 (bases 1 to 1080)
AUTHORS	Lawrence, S.M., Huddleston, K.A., Pitts, L.R., Nguyen, N., Lee, Y.C., Vann, W.F., Coleman, T.A. and Betenbaugh, M.J.
TITLE	Direct Submission
JOURNAL	Submitted (18-APR-2000) Chemical Engineering, Johns Hopkins University, 3400 N. Charles St., Baltimore, MD 21218, USA
FEATURES	Location/Qualifiers
source	1..1080
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/tissue_type="liver"
	1..1080
	/note="2-keto-3-deoxy-D-glycero-D-galacto-nononic acid biosynthetic ability"
	/codon_start=1
	/product="N-acetylneuraminic acid phosphate synthase"
	/protein_id="AAF5261.1"
	/db_xref="GI:8453156"
	/translation="MPLELEPCGRVWGQHPCFIABIGNHQDLDVAKRMIRMAK"
CDS	



Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 9 Row: d Column: 7  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 12056472.

## FEATURES source

source	gene	CDS	misc_feature
1..1181	/organism="Homo sapiens"	/product="N-acetylneuraminic acid phosphate synthase"	
/mol_type="mRNA"	/note="NANS"	/codon_start=1	
/db_xref="taxon:9606"	/note="synonym: SAS"	/protein_id="AAH19315.1"	
/clone="MGC:4339 IMAGE:2821389"	/db_xref="LocusID:54187"	/db_xref="GI:17939512"	
/tissue_type="Lung, small cell carcinoma"	/db_xref="MIM:605202"	/db_xref="LocusID:54187"	
/clone_lib="NH MGC.7"			
/lab_host="DH10B-R"			
/note="vector: pOTB7"	1..1181		
	/gene="NANS"		
	/note="synonym: SAS"		
	/db_xref="LocusID:54187"		
	/db_xref="MIM:605202"		
	61..1140		
	/codon_start=1		
	/product="N-acetylneuraminic acid phosphate synthase"		
	/protein_id="AAH19315.1"		
	/db_xref="GI:17939512"		
	/db_xref="LocusID:54187"		
	/translation="MPLELELCGRWVGQHCPFIABIGQNHQGLDVAKNRMIRAK ECGADCAKFKQSELEKFLFNKRALDRPTXSHWKGKTYGEHRKLFESHQYRELQRYA BEVGIFFTASQMDVMEVPELHNLNPFYKVGSGDTNNFYLEKTAAGRMVTSQMGQ SMDTMKVQYQVQVLEPNFPCFOCTSAYPLOQEDVNLRLVISEYQLFPDIPICYSGE TGIAISVAALGAALGLERHITLDKTMKGDHSASLEPGEALAEVLSRLVERALGSP TKLILPEMCAANEKLGKSVAKVVKIPGTTILTMDLTLVKVGEFGKGPFPDIFNLVNGK VLTVREDDTIMEELGNHNGKKTKS"		
	82..1095		
	/note="Spss; Region: Sialic acid synthase [cell envelope biogenesis, outer membrane]"		
	/db_xref="rnn.COC20ae"		

## ORIGIN

Query Match	99.7%	Score 1076.8	DB 9	Length 1181
Best Local Similarity	99.8%	Pred. No. 1.4e-255		
Matches 1078	Conservative 0	Mismatches 2	Indels 0	Gaps 0
QY	1	ATCCCGCTGAGCTGGAGCTGTGTCCCGGGCGCTGGTGGGCGGGCAACACCCGTGCTTC	60	
Db	61	ATCCCGCTGAGCTGGAGCTGTGTCCCGGGCGCTGGTGGGCGGGCAACACCCGTGCTTC	120	
QY	61	ATCATTTGCCGAGATCGGCCCAAGAACCCACAGGGCGACCTGGACGTAGCCAAAGCGCATGATC	120	
Db	121	ATCATTTGCCGAGATCGGCCCAAGAACCCACAGGGCGACCTGGACGTAGCCAAAGCGCATGATC	180	
QY	121	CGCATGCCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCAGAAAGTAGTGAGCTAGAAATTC	180	
Db	181	CGCATGCCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCAGAAAGTAGTGAGCTAGAAATTC	240	
QY	181	AAGTTTAAATCGGAAAGCCTTGGAGAGGCCATACACTCGAAGCATTTCTCTGGGGGAAGACG	240	
Db	241	AAGTTTAAATCGGAAAGCCTTGGAGAGGCCATACACTCGAAGCATTTCTCTGGGGGAAGACG	300	
QY	241	TACGGGGAGACAAAACGACATCTGGAGTTCAGGCCATGACCCAGTACAGGGAGCTGCACAGG	300	
Db	301	TACGGGGAGACAAAACGACATCTGGAGTTCAGGCCATGACCCAGTACAGGGAGCTGCACAGG	360	
QY	301	TACGCCGAGAGGTTGGGATCTCTTCACTGCCTCTGGCATGGATGGAGATGGCAGTTGAA	360	
Db	361	TACGCCGAGAGGTTGGGATCTCTTCACTGCCTCTGGCATGGATGGAGATGGCAGTTGAA	420	
QY	361	TTCTCTGCATGAACCTGAATGTTCCCAATTTTTCAAAAGTTGGATCTCGAGACACTAATAATTTT	420	
Db	421	TTCTCTGCATGAACCTGAATGTTCCCAATTTTTCAAAAGTTGGATCTCGAGACACTAATAATTTT	480	
QY	421	CCATTATCTGAAAAGACAGGCCAAAAAAGGTGCCCAATGGTGTATCTCTCAAGTGGGATCGAG	480	

Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 22388257  
 MEDLINE  
 12477932  
 PUBLISHED  
 REFERENCE  
 2 (bases 1 to 1229)  
 Strausberg, R.  
 Direct Submission  
 Submitted (03-NOV-2000) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 On Aug 19, 2003 this sequence version replaced gi:12652538.  
 Contact: MGC help desk  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
 contact: [amadansystemsbiology.org](mailto:amadansystemsbiology.org)  
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha  
 Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IMAGE Plate: 7 Row: f Column: 1  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 12056472.  
 Location/Qualifiers  
 1..1229  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:934 IMAGE:3505324"  
 /tissue type="Placenta, choriocarcinoma"  
 /clone\_lib="NIH MGC\_21"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOTB7"  
 1..1229  
 /gene="NANS"  
 /note="synonym: SAS"  
 /db\_xref="LocusID:54187"  
 /db\_xref="MIM:605202"  
 79..1158  
 /codon\_start=1  
 /product="N-acetylneuraminic acid phosphate synthase"  
 /protein\_id="AAH00008.1"  
 /db\_xref="GI:12652539"  
 /db\_xref="LocusID:54187"  
 /translation="MPLELEUCGRWVGQHPCFIIEIGQNHQGDLDVAKRMIRMAK  
 ECGADKAFQSELEFNRKALRPYTSKHSWGKTYGEHRHLEFSDQVREIQRVA  
 SEVDGFTTASMDVMEFVHELNVPPFKVGSDTNPNPYLEKAKGRPMVISGMO  
 SMPTMKVOYQIVPLNPNFCLOCTSAVPLQPEDVNLVISEYOKLPIDIPIGYSGH  
 TGLISVAALVALGAKLERHITLDKTHKSGDSHASLEPGEALVRSVRLVERALGSP  
 TKQLLPEMACNEKLGSVVAKVKIPEGITLLMDLIVKVGEPKGYPPEDIFNLVGGK  
 VLVTVEEDDTISLELVNHHGKKIS"  
 100..1113  
 /note="Sp52; Region: Sialic acid synthase [Cell envelope  
 biogenesis, outer membrane]"  
 /db\_xref="CDD:C062089"  
 misc\_feature  
 Query Match 99.7%; Score 1076.8; DB 9; Length 1229;  
 Best Local Similarity 99.8%; Pred. No. 1.4e-255;  
 Matches 1078; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 ORIGIN  
 1 ATGCGCTGGAGCTGGAGCTGTGCCCGGGCGCTGGGTGGGGCGGCAACACCCGTGCTTC 60  
 79 ATGCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGGCGGCAACACCCGTGCTTC 138

QY 61 ATCATTGCGGAGATCGGCCAGAACCCACGAGGGGACCTTGGAGCTAGCCAGCGCATGATC 120  
 Db 139 ATCATTGCGGAGATCGGCCAGAACCCACGAGGGGACCTTGGAGCTAGCCAGCGCATGATC 198  
 QY 121 CGCATGGCCAAAGGAGTGTGGGGCTGATTGTGCCAAGTTTCCAGAAAGAGTGAGCTGAATTC 180  
 Db 199 CGCATGGCCAAAGGAGTGTGGGGCTGATTGTGCCAAGTTTCCAGAAAGAGTGAGCTGAATTC 258  
 QY 181 AAGTTTAATCGAAAGCCTTGGAGAGGCCATACACCTCGAAAGCAATTCCTGGGGGAGAGCG 240  
 Db 259 AAGTTTAATCGAAAGCCTTGGAGAGGCCATACACCTCGAAAGCAATTCCTGGGGGAGAGCG 318  
 QY 241 TACGGGGAGCAACACGACATCTGGAGTTTCAGCCATGACCCAGTACAGGAGCTGCAGAGG 300  
 Db 319 TACGGGGAGCAACACGACATCTGGAGTTTCAGCCATGACCCAGTACAGGAGCTGCAGAGG 378  
 QY 301 TACGCCGAGGAGTTGGGATCTTCTTCACTGCTCTGGCATGGATGAGATGGCAGTTGAA 360  
 Db 379 TACGCCGAGGAGTTGGGATCTTCTTCACTGCTCTGGCATGGATGAGATGGCAGTTGAA 438  
 QY 361 TTCCTGCAATGAACCTGAATGTTCCATTTTCAAAAGTTGGATCTGGAGACACTAATAATTTT 420  
 Db 439 TTCCTGCAATGAACCTGAATGTTCCATTTTCAAAAGTTGGATCTGGAGACACTAATAATTTT 498  
 QY 421 CCTTATCTGGAAGACAGCCAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGAG 480  
 Db 499 CCTTATCTGGAAGACAGCCAAAGGTCGCCCAATGGTGATCTCCAGTGGGATGAG 558  
 QY 481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTGCG 540  
 Db 559 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTGCG 618  
 QY 541 TTCCTGCAATGATACAGCGCATACCCGCTCAGCGCTGAGAGCGTCAACCTCGGGTCTATC 600  
 Db 619 TTCCTGCAATGATACAGCGCATACCCGCTCAGCGCTGAGAGCGTCAACCTCGGGTCTATC 678  
 QY 601 TCGGAATATCAGAGCTCTTCTCGACATTCGCCATAGGTTATCTGGGCATGAAACAGGC 660  
 Db 679 TCGGAATATCAGAGCTCTTCTCGACATTCGCCATAGGTTATCTGGGCATGAAACAGGC 738  
 QY 661 ATAGCGATATCTGTGGCGCGAGTGGCTCTGGGGGCGAAGGTGTTGGAACGTCACATAACT 720  
 Db 739 ATAGCGATATCTGTGGCGCGAGTGGCTCTGGGGGCGAAGGTGTTGGAACGTCACATAACT 798  
 QY 721 TTGACAAAGACCTGGAAGGGGAGTGACCACTCGGCGCTCGCTGGAGCCTCGGAGAACTGGCC 780  
 Db 799 TTGACAAAGACCTGGAAGGGGAGTGACCACTCGGCGCTCGCTGGAGCCTCGGAGAACTGGCC 858  
 QY 781 GAGCTGGTGGCTCAGTGGCTGTTGGAGGGTGGCTGGGGCTCCCAACCAAGCGAGCTG 840  
 Db 859 GAGCTGGTGGCTCAGTGGCTGTTGGAGGGTGGCTGGGGCTCCCAACCAAGCGAGCTG 918  
 QY 841 CTGCCCTGTGAGATGGCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAG 900  
 Db 919 CTGCCCTGTGAGATGGCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAG 978  
 QY 901 ATTCGGGAAGGCACTTCTTAACTAATGACATGCTCACCGTGAAGGTGGGTGAGCCCAA 960  
 Db 979 ATTCGGGAAGGCACTTCTTAACTAATGACATGCTCACCGTGAAGGTGGGTGAGCCCAA 1038  
 QY 961 GCCTATCTCTGAAGACATCTTAACTAGTGGGCAAGAGGTCTCTGGTCACTGTTGAA 1020  
 Db 1039 GGCTATCTCTGAAGACATCTTAACTAGTGGGCAAGAGGTCTCTGGTCACTGTTGAA 1098  
 QY 1021 GAGGATGACACCATCATGGAAGAAATGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080  
 Db 1099 GAGGATGACACCATCATGGAAGAAATGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1158

RESULT 5  
 AR339544  
 LOCUS  
 DEFINITION Sequence 1035 from patent US 656962.  
 AR339544 1230 bp DNA linear PAT 17-AUG-2003

```
ACCESSION AR339544
VERSION AR339544.1 GI:33726401
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
  1 (bases 1 to 1230)
  Tang,Y.F., Zhou,P. and Drmanac,R.T.
  Nucleic acids and polypeptides
  TITLE Patent: US 6569662-A 1035 27-MAY-2003;
  JOURNAL Location/Qualifiers
  FEATURES
    source
      1..1230
        /organism="unknown"
        /mol_type="genomic DNA"
  ORIGIN
    Query Match          99.7%; Score 1076.8; DB 6; Length 1230;
    Best Local Similarity 99.8%; Pred. No. 1.4e-255;
    Matches 1078; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGCGCTGGAGCTGAGCTGTGTCCGGGCGCTGGTGGGGCGGCAACACCCGTCCTTC 60
Db 119 ATGCGCTGGAGCTGAGCTGTGTCCGGGCGCTGGTGGGGCGGCAACACCCGTCCTTC 178
QY 61 ATCATTCGCGAGATCGCCAGAACACACAGGGCGACCTGGAGTAGCCAGGCGCATGATC 120
Db 179 ATCATTCGCGAGATCGCCAGAACACACAGGGCGACCTGGAGTAGCCAGGCGCATGATC 238
QY 121 CGCATGGCCAGAGGTGTGGGCTGATTGTGCGCAAGTTCGAAAGAGTGAGTAGAATTC 180
Db 239 CGCATGGCCAGAGGTGTGGGCTGATTGTGCGCAAGTTCGAAAGAGTGAGTAGAATTC 298
QY 181 AGTTTAAATCGAAGCCCTTGAGAGGCGCATACACTCGAAGCATTCCTGGGGAGAGCG 240
Db 299 AGTTTAAATCGAAGCCCTTGAGAGGCGCATACACTCGAAGCATTCCTGGGGAGAGCG 358
QY 241 TACGGGGAGCACAAACGACATCTGGAGTTTCAGCATGACACAGTACAGGAGCTGCAGAGG 300
Db 359 TACGGGGAGCACAAACGACATCTGGAGTTTCAGCATGACACAGTACAGGAGCTGCAGAGG 418
QY 301 TACGGCGAGGAGTTGGGATCTTTCACTGCTCTCGGCATGGATGAGATGGCAGTTGAA 360
Db 419 TACGGCGAGGAGTTGGGATCTTTCACTGCTCTCGGCATGGATGAGATGGCAGTTGAA 478
QY 361 TTCTTCGATGAACGAAATGTTCCATTTTCAAGTTGGATCTGGAGACACTAATATTTT 420
Db 479 TTCTTCGATGAACGAAATGTTCCATTTTCAAGTTGGATCTGGAGACACTAATATTTT 538
QY 421 CTTTATCTGGAAAAGACAGCCAAAAGGTGCGCCCAATGGTGATCTCCAGTGGGATGCAG 480
Db 539 CTTTATCTGGAAAAGACAGCCAAAAGGTGCGCCCAATGGTGATCTCCAGTGGGATGCAG 598
QY 481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTGC 540
Db 599 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTGC 658
QY 541 TTCTTCGATGTACAGCGCATACCGCTCCAGCTCGAGCATGAGACGTCAACTGCGGTCATC 600
Db 659 TTCTTCGATGTACAGCGCATACCGCTCCAGCTCGAGCATGAGACGTCAACTGCGGTCATC 718
QY 601 TCGGAATATCAGAACTCTTTCTGACATTCCTAGAGGTATCTTGGGATGAAGACAGC 660
Db 719 TCGGAATATCAGAACTCTTTCTGACATTCCTAGAGGTATCTTGGGATGAAGACAGC 778
QY 661 ATAGCGATATCTGTGGCGCAGTGGCTCTGGGGGCAAGGTGTGGAACGTCACTAATCT 720
Db 779 ATAGCGATATCTGTGGCGCAGTGGCTCTGGGGGCAAGGTGTGGAACGTCACTAATCT 838
QY 721 TTGGACAGACCTGAAGGGAGTGACCACTCGGCTCGCTCGAGCTGGAGACTGGCC 780
Db 839 TTGGACAGACCTGAAGGGAGTGACCACTCGGCTCGCTCGAGCTGGAGACTGGCC 898
QY 781 GAGCTGGTTCAGTGCCTTCTGTGGAGCGTGCCTCGGCTCCCCCAACCAAGCAGCTG 840
```

```
Db 899 GAGCTGGTGGCTAGTGGCTTTGTGAGCGCTGCCCTGGGCTCCCAACCAAGCAGCTG 958
QY 841 CTGCGCTCTGAGATGGCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGA 900
Db 959 CTGCGCTCTGAGATGGCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGA 1018
QY 901 ATTCCGGAAGGACCAATCTTAACAATGACATGCTCACCGTGAAGGTGGGTGAGCCAAA 960
Db 1019 ATTCCGGAAGGACCAATCTTAACAATGACATGCTCACCGTGAAGGTGGGTGAGCCAAA 1078
QY 961 GCCTATCTCTTGAAGACATCTTTAATCTAGTGGGCAAGAGTCTCTGTCTCACTGTGAA 1020
Db 1079 GCCTATCTCTTGAAGACATCTTTAATCTAGTGGGCAAGAGTCTCTGTCTCACTGTGAA 1138
QY 1021 GAGGATGACACCATCATGGAAGATTGGTAGAATAATCATGGCAAAAAATCAAGTCTTAA 1080
Db 1139 GAGGATGACACCATCATGGAAGATTGGTAGAATAATCATGGCAAAAAATCAAGTCTTAA 1198

RESULT 6
AX877224 1170 bp DNA linear PAT 17-DEC-2003
LOCUS Sequence 12129 from Patent EP1074617.
DEFINITION AX877224
ACCESSION AX877224
VERSION AX877224.1 GI:40031960
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1
  Ota,T., Isegai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
  Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  Primers for synthesising full-length cDNA and their use
  TITLE Patent: EP 1074617-A 12129 07-FEB-2001;
  JOURNAL Research Association for Biotechnology (JP)
  FEATURES
    source
      1..1170
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
        71..1150
          /note="unnamed protein product"
          /codon_start=1
          /protein_id="CAE89627.1"
          /db_xref="GI:40031961"
          /translation="MPLLELCPRWVGQHPFCFIIAETIGNHQDLDVAKERMIRMAK
          ECGADCAKFKSELEFKFNKALDRPYTKSHWGKTYGHEKRLHLEFSDQVRELQRYA
          BEVGITFTASGMDMAVEFLHENVPPFKVSGSDTNPNPYLEKTKAKGRPMVTSQMQ
          SMDTWKQVQIVKPLNPNFCLOCTSAYPLOPEDVNLVISEYQKLFDPIDPLIGYSGHE
          TGAISVAALVKVLEKTLTDKTKWGSDBSLSEPEGEIAELVRSVRLVBRALGSP
          TKQLLPCMACNEKLGKSVAKVKIPTGILTMDMLTVKVGEPKGPYPEDIPNLVGRK
          VLVTVEEDDTIMEELVDNHGKKIKS"
    CDS
      1
      /organism="Homo sapiens"
      /mol_type="unassigned DNA"
      /db_xref="taxon:9606"
      71..1150
        /note="unnamed protein product"
        /codon_start=1
        /protein_id="CAE89627.1"
        /db_xref="GI:40031961"
        /translation="MPLLELCPRWVGQHPFCFIIAETIGNHQDLDVAKERMIRMAK
        ECGADCAKFKSELEFKFNKALDRPYTKSHWGKTYGHEKRLHLEFSDQVRELQRYA
        BEVGITFTASGMDMAVEFLHENVPPFKVSGSDTNPNPYLEKTKAKGRPMVTSQMQ
        SMDTWKQVQIVKPLNPNFCLOCTSAYPLOPEDVNLVISEYQKLFDPIDPLIGYSGHE
        TGAISVAALVKVLEKTLTDKTKWGSDBSLSEPEGEIAELVRSVRLVBRALGSP
        TKQLLPCMACNEKLGKSVAKVKIPTGILTMDMLTVKVGEPKGPYPEDIPNLVGRK
        VLVTVEEDDTIMEELVDNHGKKIKS"
    ORIGIN
      Query Match          99.6%; Score 1075.2; DB 6; Length 1170;
      Best Local Similarity 99.7%; Pred. No. 3.4e-255;
      Matches 1077; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGCGCTGGAGCTGAGCTGTGTCCGGGCGCTGGTGGGGCGGCAACACCCGTCCTTC 60
Db 71 ATGCGCTGGAGCTGAGCTGTGTCCGGGCGCTGGTGGGGCGGCAACACCCGTCCTTC 130
QY 61 ATCATTCGCGAGATCGCCAGAACCAACAGGGCGACCTGGAGTAGCCAGGCGCATGATC 120
Db 131 ATCATTCGCGAGATCGCCAGAACCAACAGGGCGACCTGGAGTAGCCAGGCGCATGATC 190
QY 121 CGCATGGCCAGAGGTGTGGGCTGATTGTGCGCAAGTTCGAAAGAGTGAGTAGAATTC 180
Db 191 CGCATGGCCAGAGGTGTGGGCTGATTGTGCGCAAGTTCGAAAGAGTGAGTAGAATTC 250
QY 181 AAGTTTAAATCGAAGCCCTTGAGAGGCGCATACACTCGAAGCATTCCTCTGGGGAGAGCG 240
```

Db 251 AAGTTTAAATCGAAAGCCTTGGACAGCCATACACCTCGAAGCATTCCTGGGGAGACG 310  
Qy 241 TAGGGGACACAAAGAGCATCTGGAGTTTACGCATACAGGAGCTGCAGAGG 300  
Db 311 TAGGGGAGCACAAGAGCATCTGGAGTTTACGCATACAGGAGCTGCAGAGG 370  
Qy 301 TAGCCGAGAGGTTGGGATCTTCTCACTGCTCTGGCATGAGATGAGATGAGTTGAA 360  
Db 371 TAGCCGAGAGGTTGGGATCTTCTCACTGCTCTGGCATGAGATGAGATGAGTTGAA 430  
Qy 361 TTCCTGCATGAACTGAATTTTCATTTTCAAAGTTTGGATCTGGAGACACTAATAATTTT 420  
Db 431 TTCCTGCATGAACTGAATTTTCATTTTCAAAGTTTGGATCTGGAGACACTAATAATTTT 490  
Qy 421 CCTTATCTGAAAAGACACGCAAAAGAGTCCGCCAATGGTATCTCCAGTGGATGAG 480  
Db 491 CCTTATCTGAAAAGACACGCAAAAGAGTCCGCCAATGGTATCTCCAGTGGATGAG 550  
Qy 481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTGC 540  
Db 551 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTGC 610  
Qy 541 TTCCTGCATGTAACAGGAGTACCCGCTCCAGCTCCAGGAGCTCAACCTCGGGGTATC 600  
Db 611 TTCCTGCATGTAACAGGAGTACCCGCTCCAGGAGCTCAACCTCGGGGTATC 670  
Qy 601 TCGGAATATCAAGAGCTCTTCTCAGATTCCTCAGATTCCTCAGATTCCTCAGATTCCT 660  
Db 671 TCGGAATATCAAGAGCTCTTCTCAGATTCCTCAGATTCCTCAGATTCCTCAGATTCCT 730  
Qy 661 ATAGCGATATCTGTGGCGGAGTGGCTCTGGGGGCAAGGTGTGGAAAGCTCACATAACT 720  
Db 731 ATAGCGATATCTGTGGCGGAGTGGCTCTGGGGGCAAGGTGTGGAAAGCTCACATAACT 790  
Qy 721 TTGACAAGACTGGAAGGAGTGAACCACTCGGCTCTGGAGCTCGGAGCTGGAACTGGCC 780  
Db 791 TTGACAAGACTGGAAGGAGTGAACCACTCGGCTCTGGAGCTCGGAGCTGGAACTGGCC 850  
Qy 781 GAGCTGTGGCTGAGTGGCTCTGGAGGAGTGGCAAGTCTGTGGTGGCCAAAGCTGAA 840  
Db 851 GAGCTGTGGCTGAGTGGCTCTGGAGGAGTGGCAAGTCTGTGGTGGCCAAAGCTGAA 910  
Qy 841 CTGGCTCTGAGATGGCTCTGCAATGAGAGTGGCAAGTCTGTGGTGGCCAAAGCTGAA 900  
Db 911 CTGGCTCTGAGATGGCTCTGCAATGAGAGTGGCAAGTCTGTGGTGGCCAAAGCTGAA 970  
Qy 901 ATTCCGGAAGGACCAATCTTAACAAATGACATGCTCAACGTTGAAGGTGGTGAAGG 960  
Db 971 ATTCCGGAAGGACCAATCTTAACAAATGACATGCTCAACGTTGAAGGTGGTGAAGG 1030  
Qy 961 GCCTATCTCTGAAAGACATCTTTAATCTAGTGGCAAGAGTCTGGTCACTGTTGAA 1020  
Db 1031 GCCTATCTCTGAAAGACATCTTTAATCTAGTGGCAAGAGTCTGGTCACTGTTGAA 1090  
Qy 1021 GAGGATGACCATCATGGAAGAAATTTGATAGATATCATGCAAAATAATCAAGTCTTAA 1080  
Db 1091 GAGGATGACCATCATGGAAGAAATTTGATAGATATCATGCAAAATAATCAAGTCTTAA 1150

RESULT 7  
BD156549  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

BD156549  
Primer for synthesizing full-length cDNA and use thereof.  
BD156549  
BD156549.1 GI:27862307  
JP 2002191363-A/11392.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1. (bases 1 to 1170)  
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,

Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.  
Primer for synthesizing full-length cDNA and use thereof  
HELIX RESEARCH INSTITUTE  
OS Homo sapiens (human)  
PN JP 2002191363-A/11392  
PD 09-JUL-2002 JP 2000280990  
PF 28-JUL-2000 JP 2000280990  
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU  
PI SAITO,  
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,  
PI KEIICHI NAGAI, TETSUJI OTSUKI  
PC  
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof FH Key  
FEATURES  
source  
1..1170  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 99.6%; Score 1075.2; DB 6; Length 1170;  
Best Local Similarity 99.7%; Pred. No. 3.4e-255;  
Matches 1077; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 ATCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGTGGGGGGGCAACACCCGTGCTTC 60  
Db 71 ATCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGTGGGGGGGCAACACCCGTGCTTC 130  
Qy 61 ATCATTTCCGAGATCGGCCAGAACCAACAGGCGGACCTGGAGTAGCAGCAAGCGCATGATC 120  
Db 131 ATCATTTCCGAGATCGGCCAGAACCAACAGGCGGACCTGGAGTAGCAGCAAGCGCATGATC 190  
Qy 121 CGCATGGCAAGGAGTGGGGCTGATTTGCCAAGTTCCAGAGAGTGAGTAGAATTC 180  
Db 191 CGCATGGCAAGGAGTGGGGCTGATTTGCCAAGTTCCAGAGAGTGAGTAGAATTC 250  
Qy 181 AAGTTTAAATCGGAAGCCTTGGAGAGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240  
Db 251 AAGTTTAAATCGGAAGCCTTGGAGAGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 310  
Qy 241 TACGGGAGCAGCAACAGACATCTCGAGTTTCAGCCATGACAGTACAGGAGCTGCAGAGG 300  
Db 311 TACGGGAGCAGCAACAGACATCTCGAGTTTCAGCCATGACAGTACAGGAGCTGCAGAGG 370  
Qy 301 TACGGGAGGAGTTGGGATCTTCTCACTGCTCTGGCATGAGATGAGATGAGTTGAA 360  
Db 371 TACGGGAGGAGTTGGGATCTTCTCACTGCTCTGGCATGAGATGAGATGAGTTGAA 430  
Qy 361 TTCCTGCATGAACTGAATTTTCAATTTTCAAAGTTTGGATCTGGAGACACTAATAATTTT 420  
Db 431 TTCCTGCATGAACTGAATTTTCAATTTTCAAAGTTTGGATCTGGAGACACTAATAATTTT 490  
Qy 421 CCTTATCTGAAAAGACACGCAAAAGAGTCCGCCAATGGTATCTCCAGTGGATGAG 480  
Db 491 CCTTATCTGAAAAGACACGCAAAAGAGTCCGCCAATGGTATCTCCAGTGGATGAG 550  
Qy 481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTGC 540  
Db 551 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTGC 610  
Qy 541 TTCCTGCATGTAACAGGAGTACCCGCTCCAGCTCCAGGAGCTCAACCTCGGGGTATC 600  
Db 611 TTCCTGCATGTAACAGGAGTACCCGCTCCAGGAGCTCAACCTCGGGGTATC 670  
Qy 601 TCGGAATATCAAGAGCTCTTCTCAGATTCCTCAGATTCCTCAGATTCCTCAGATTCCT 660  
Db 671 TCGGAATATCAAGAGCTCTTCTCAGATTCCTCAGATTCCTCAGATTCCTCAGATTCCT 730



```

QY 661 ATAGCGATATCTGTGGCGCAGTGGCTCTGTGGGGCCAAAGTGTGGAACGTCACATAACT 720
DB 731 ATAGCGATATCTGTGGCGCAGTGGCTCTGTGGGGCCAAAGTGTGGAACGTCACATAACT 790
QY 721 TTGGACAGAGCTGGAAGGGAGTGACCACTCGGCTCGCTGGAGCCCTGGAGAACTGGCC 780
DB 791 TTGGACAGAGCTGGAAGGGAGTGACCACTCGGCTCGCTGGAGCCCTGGAGAACTGGCC 850
QY 781 GAGCTGGTGGCTGAGTGGCTCTGTGGAGCGTGGCTGGGCTCCCAACCAAGCAGCTG 840
DB 851 GAGCTGGTGGCTGAGTGGCTCTGTGGAGCGTGGCTGGGCTCCCAACCAAGCAGCTG 910
QY 841 CTGCGCTGTGAGATGGCTGCAATGAGAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
DB 911 CTGCGCTGTGAGATGGCTGCAATGAGAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 970
QY 901 ATTCCGGAGGACCACTTACCAATGACATGCTCACCGTGAAGTGGTGGAGCCAAA 960
DB 971 ATTCCGGAGGACCACTTACCAATGACATGCTCACCGTGAAGTGGTGGAGCCAAA 1030
QY 961 GCGTATCTCTGTAAGACATCTTTAATCTAGTGGGCAAGAGTCTGTGTCATCTGTGAA 1020
DB 1031 GCGTATCTCTGTAAGACATCTTTAATCTAGTGGGCAAGAGTCTGTGTCATCTGTGAA 1090
QY 1021 GAGGATGACACCATCATGGAAGAATTTGAGATATCATGCGCAAAAAAATCAAGCTTAA 1080
DB 1091 GAGGATGACACCATCATGGAAGAATTTGAGATATCATGCGCAAAAAAATCAAGCTTAA 1150

RESULT 8
AK001659
LOCUS
DEFINITION
  Homo sapiens cDNA FLJ10797 fis, clone NT2RP4000657, weakly similar
  to SPORE COAT POLYSACCHARIDE BIOSYNTHESIS PROTEIN SPSE.
ACCESSION
  AK001659
VERSION
  1
KEYWORDS
  oligo capping; fis (full insert sequence).
SOURCE
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1
  Isogai,T., Ota,T., Hayaashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
  Nishikawa,T., Nagai,K., Sugano,S., Aotsuka,S., Yoshikawa,Y.,
  Matsumawa,H., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,
  Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuho,Y. and Sasaki,N.
  NEDO human cDNA sequencing project
  Unpublished
  2 (bases 1 to 1170)
  Isogai,T. and Otsuki,T.
  Direct Submission
  Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,
  Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
  (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
  NEDO human cDNA sequencing project supported by Ministry of
  International Trade and Industry of Japan; cDNA full insert
  sequencing; Research Association for Biotechnology; cDNA library
  construction, 5'- & 3'-end one pass sequencing and clone selection;
  Helix Research Institute (supported by Japan Key Technology Center
  etc.) and Department of Virology, Institute of Medical Science,
  University of Tokyo.
FEATURES
  Location/Qualifiers
  1..1170
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="NT2RP4000657"
    /cell_line="NT2"
    /cell_type="teratocarcinoma"
    /clone_lib="NT2RP4"
    /notes="cloning vector: pME18SFL3-mRNA from NT2 neuronal
    precursor cells after 2-weeks retinoic acid (RA)
  
```

```

induction."
71..1150
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA91818.1"
/db_xref="GI:7023053"
/translation="MPLLELCPRVWGQHPFIIARIQNHQGLDVAKRMIRMAK
ECGADCAKFKSELEFKENRKALDRPVTSKHMGKTYGEHKLHLEFSDVRELORYA
EVGILFFASGDMENAVEFLHNLNVPFFKVGSGDTNNPYLEKTAKKGRPVWISGMQ
SMDTKQYQIVKPLNFCFLQKTSAYPLQPEDVNLKRVISYIYQKLFDPILPISGHE
TGIAISVAALVQITKVLERTHTLDKTKWGSDBSLSEPELAEILRVRLVERALGSP
TKQLLPCEMACNEKLGKSVAKVKIPEGTILTMDLTKVKGEPKGPYPEDIPNLVGGK
VLVTVEEDDTIMEELVDNHNKGIKS"

ORIGIN
Query Match      99.6%; Score 1075.2; DB 9; Length 1170;
Best Local Similarity 99.7%; Pred. No. 3.4e-255;
Matches 1077; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGCCGCTGGAGCTGAGCTGTCCCGGCGCTGGTGGGGGGGCAACACCCGCTGCTTC 60
DB 71 ATGCCGCTGGAGCTGAGCTGTCCCGGCGCTGGTGGGGGGGCAACACCCGCTGCTTC 130
QY 61 ATCATTTGCCGAGATCGGCCAGAACCAACAGGGCGACCTGGACCTAGCCAGCGCATGATC 120
DB 131 ATCATTTGCCGAGATCGGCCAGAACCAACAGGGCGACCTGGACCTAGCCAGCGCATGATC 190
QY 121 CGCATGGCCAGAGGTGTGGGCGCTGATTTGTGCCAAGTTCCAGAAAGAGTAGTAGAATTC 180
DB 191 CGCATGGCCAGAGGTGTGGGCGCTGATTTGTGCCAAGTTCCAGAAAGAGTAGTAGAATTC 250
QY 181 AGTTTAAATCGAAAGACCTTGGAGAGGCCATACACTCGAAGACATTCCTGGGGGAAGACG 240
DB 251 AGTTTAAATCGAAAGACCTTGGAGAGGCCATACACTCGAAGACATTCCTGGGGGAAGACG 310
QY 241 TACGGGGAGCACAAACAGCATCTGGAGTTTACGCCATGACACAGGAGTGCAGAGG 300
DB 311 TACGGGGAGCACAAACAGCATCTGGAGTTTACGCCATGACACAGGAGTGCAGAGG 370
QY 301 TACGGGGAGGAGTTGGGATCTTCTTCACTGCTCTGGCATGGATGAGATGCGAGTTGAA 360
DB 371 TACGGGGAGGAGTTGGGATCTTCTTCACTGCTCTGGCATGGATGAGATGCGAGTTGAA 430
QY 361 TTCTTGCATGAATCTGAATGTTTCCATTTTCAAAGTTGGATCTGGAGACATAATAATTTT 420
DB 431 TTCTTGCATGAATCTGAATGTTTCCATTTTCAAAGTTGGATCTGGAGACATAATAATTTT 490
QY 421 CCTTATCTGAAAAAGACAGCCCAAAAAGGTCCGCCATGAGTCTCCAGTGGGATGCGAG 480
DB 491 CCTTATCTGAAAAAGACAGCCCAAAAAGGTCCGCCATGAGTCTCCAGTGGGATGCGAG 550
QY 481 TCAATGGACACCATGAGCAAGTTTATTCAGATCGTGAAGCCCTCAACCCCACTTCTGC 540
DB 551 TCAATGGACACCATGAGCAAGTTTATTCAGATCGTGAAGCCCTCAACCCCACTTCTGC 610
QY 541 TTCTTGCAGTGTACAGCGCATACCGCTCCAGCCTGAGGAGCTCAACCTGGGGTCAATC 600
DB 611 TTCTTGCAGTGTACAGCGCATACCGCTCCAGCCTGAGGAGCTCAACCTGGGGTCAATC 670
QY 601 TCGGAATATCAGAAGCTCTTTCTGACATTCCTCATAGGTATTTCTGGGATGAAACAGGC 660
DB 671 TCGGAATATCAGAAGCTCTTTCTGACATTCCTCATAGGTATTTCTGGGATGAAACAGGC 730
QY 661 ATAGCGATATCTGTGGCGCAGTGGCTCTGGGGGGCCAAAGTGTGGAACGTCACATAACT 720
DB 731 ATAGCGATATCTGTGGCGCAGTGGCTCTGGGGGGCCAAAGTGTGGAACGTCACATAACT 790
QY 721 TTGGACAGAGCTGGAAGGGAGTGACCACTCGGCTCGCTGGAGCCCTGGAGAACTGGCC 780
DB 791 TTGGACAGAGCTGGAAGGGAGTGACCACTCGGCTCGCTGGAGCCCTGGAGAACTGGCC 850
QY 781 GAGCTGGTGGCTGAGTGGCTCTGTGGAGCGTGGCTGGGCTCCCAACCAAGCAGCTG 840
  
```



```

Db      851 GAGCTGGTGGCGTCACTGGCTCTGTGGAGCGTGCCTGGGCTCCCAACCAAGCAGCTG 910
QY      841 CTGCCCTGTGAGATGCCCTGCAATGAGAAGCTGGGCAAGCTCTGTGGTGCCCAAGTGAAA 900
Db      911 CTGCCCTGTGAGATGCCCTGCAATGAGAAGCTGGGCAAGCTCTGTGGTGCCCAAGTGAAA 970
QY      901 ATTCCGGAAGGACCACTTCAACAATGGACATGCTCACCGTCAAGTGGTGGTGGAGCCCAA 960
Db      971 ATTCCGGAAGGACCACTTCAACAATGGACATGCTCACCGTCAAGTGGTGGTGGAGCCCAA 1030
QY      961 GCCTATCTCTCTGAAGACATCTTTAATCTAGTGGGCAAGAAGTCTCTGTCTACTGTGAA 1020
Db      1031 GCCTATCTCTCTGAAGACATCTTTAATCTAGTGGGCAAGAAGTCTCTGTCTACTGTGAA 1090
QY      1021 GAGGATGACACCATCATGGAAGAATTGTAGTATATCATGCGCAAAAATAATCAAGTCTTAA 1080
Db      1091 GAGGATGACACCATCATGGAAGAATTGTAGTATATCATGCGCAAAAATAATCAAGTCTTAA 1150

RESULT 9
LOCUS   BC003307
DEFINITION Mus musculus N-acetylneuraminic acid synthase (sialic acid
ACCESSION BC003307
VERSION   BC003307.1 GI:13097041
KEYWORDS MGC.
SOURCE   Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 1247)
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schneitz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, D.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettner, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 1247)
Straussberg, R.
Direct Submission
Submitted (20-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hqsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu

```

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louleseg, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 6 Row: 1 Column: 3  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 16716466.

## FEATURES

```

Location/Qualifiers
1..1247
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129,C57BL/6J,FVB/N"
/db_xref="taxon:10090"
/clone="MGC:6982 IMAGE:3154822"
/tissue_type="Mammary tumor. Brca1-/fl; MMTV-Cre model. 10 months old, gross tissue."
/clone_lib="NCI CGAP_Mam3"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
1..1247
/gene="Nans"
/db_xref="LocusID:94181"
/db_xref="MGI:2149820"
124..1203
/codon_start=1
/product="N-acetylneuraminic acid synthase (sialic acid synthase)"
/protein_id="AAH03307.1"
/db_xref="GI:13097042"
/db_xref="LocusID:94181"
/translation="MPELELCPRWVGKHPFIIAIGONHOGIDIVAKRMIRTAKECAGADCFKQSELEFKNKALRPVTSKSHGKTEGHEKHLFESHDQYKELQSYA QEIGIFRTASQMDSEMAVEFLHNLNVPFKVSGGDTNNPFLYKTAKRPMVSSGMQ SMDPMQYQIVKPLNPNFCLQCTSAIPQLPEDANLKVISEYKLFDPIDPIGSGHETGIAISVAALGAKVLRHITLTKWGSDSASLEPGEALAEIVRSVLRVLRALGSP TKQLLPCEMACNEKLGKSVAKVPIPAQTTLTLDMLTVKVGEPKGYPPEDIFNLGKK VLVTIEEDTVMEESVESHKKIKA"
145..1176
/note="Spss; Region: Sialic acid synthase [Cell envelope biogenesis, outer membrane]"
/db_xref="CDD:COG2089"

```

## misc\_feature

## ORIGIN

```

Query Match      81.8%; Score 883.2; DB 10; Length 1247;
Best Local Similarity 88.6%; Pred. No. 1.2e-207;
Matches 957; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 1 ATCCGCTGGAGCTGGAGCTGTGTCCTGGGCGCTGGTGGCGGGCAACACCCGTGCTTC 60
Db 124 ATCCGCTGGAACTGGAGCTGTGTCCTGGGCGCTGGTGGCGGGCAACACCCGTGCTTC 183
QY 61 ATCATTCCGAGATCGCCGCAACACCGGCGACCTGGACGTAGCCAGCGCATGATC 120
Db 184 ATCATCGGAGATCGGCCAGAACCAACCAAGGACATAGATGTGGCCAGCGCATGATC 243
QY 121 CGCATGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGGAGTAGAATTC 180
Db 244 CGCACTGCCAAGGAGTGTGGGGCGCATGCTGGCTAAGTTTCAGAAGAGCGAGTGGAGTTC 303
QY 181 AAGTTTAAATCGGAAAGCTTTGGAGAGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
Db 304 AAGTTTAAACGGAAGCGCTGGAGAGACCATATATCTTCGAAGCATTCATGGGGGAAGACG 363
QY 241 TAGCGGAGACAAACGACATCTGGAGTTTCCAGCATGACCATGACAGGAGTGCAGAGG 300
Db 364 TATGGGAGACAAAGCGGCATCTGGAATTCAGCCACGACCATGACAGGAGTGCAGAGC 423
QY 301 TAGCCCGAGGAGTTGGGATCTTCTTCACATGCTCTGCGCATGGATGAGATGGAGTTGAA 360
Db 424 TATGGCAGGAGATCGGCACTTCTTCTTCTGCGCATGGACGAGATGGCAGTTGAG 483

```

Qy 361 TTCCTGATGAACTGAATGTTCCATTTTTCATTTTCAAGTTGGATCTGGAGACACTAATATTTT 420  
Db 484 TTTCTGCACGAACGAATGTTCCCTTTTTCATTTTCAAGTTGGATCTGGGACACTAACAATTTT 543  
Qy 421 CTTTATCTGGAAGAACAGACGCAAAAAGTGCCTCAATGGTGTATCTCCAGTGGGATGCAG 480  
Db 544 CCTTACCTGGAAGAACAGACGCAAAAAGTGCCTCAATGGTGTATCTCCAGTGGGATGCAG 603  
Qy 481 TCAATGGAACCATGAAAGCAAGTTTATCAGATCTGGAAGCCCTCAACCCCAACTTCTGC 540  
Db 604 TCAATGGAACCATGAAAGCAAGTTTATCAGATCTGGAAGCCCTCAACCCCAACTTCTGC 663  
Qy 541 TTTCTGCACTGACAGGCAATACCGCTCCAGCGCTGAGGAGTCAACCTCGCGGTCTATC 600  
Db 684 TTTCTTCCATGACAGGCGGTACCACTACAGCCGAGGATGCCAATCTGGCGGTCTATC 723  
Qy 601 TCGGAATATCAGAACTCTTTCTCTGACATTTCCATAGGTTATTTCTGGGATGAACAGGC 660  
Db 724 TCGGAATATCAGAACTCTTTCTCTGACATTTCCATAGGTTATTTCTGGGATGAACAGGC 783  
Qy 661 ATAGGATATCTGTGGCGAGTGTCTGTGGGGGCAAGGTTTGGAGCTCACATACT 720  
Db 784 ATGCGCATATCTGTGGCGAGTGTCTGTGGGGGCAAGGTTTGGAGCTCACATACT 843  
Qy 721 TTGGACAAGACCTGGAAGGGAGTGACCACTCGGCTCTGCTGAGCTGGAGAACTGGCC 780  
Db 844 TTGGACAAGACCTGGAAGGGAGTGACCACTCGGCTCTGCTGAGCTGGAGAACTGGCC 903  
Qy 781 GAGCTGTGTGGCTGAGTGGCTTTTGTGGAGCGTGGCTGGGCTCCCCCAACGAGCTG 840  
Db 904 GAGCTGTGTGGCTGAGTGGCTTTTGTGGAGCGTGGGCTCCCCCAACGAGCTG 963  
Qy 841 CTGCGCTGTGAGATGGCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAA 900  
Db 964 CTGCGCTGTGAGATGGCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAA 1023  
Qy 901 ATTCCGGAAGGACCATTTCAATGAGATGACATGCTCACGCTGAAGTGGGTGAGCCCAA 960  
Db 1024 ATCCGAGAGGACCACTGACCTGACCTGGACATGCTCACTGTAAGTGGGGAGGCCAA 1083  
Qy 961 GCGTATCTCTCTGAAGACATCTTTAATCTAGTGGGCAAGAGGTCTGTGTCACTGTGAA 1020  
Db 1084 GCGTATCTCTCTGAAGACATCTTTAATCTAGTGGGCAAGAGGTCTGTGTCACTGTGAA 1143  
Qy 1021 GAGGATGACACCATCATGGAAGAAATGTAGATATATCATGGCAAAAATCAAGCTTAA 1080  
Db 1144 GAAGATGACACCGTCTATGGAAGAAATCCGTGGAAAGTCAAGCAAGAAATCAAGGCTTAA 1203

RESULT 10  
BC057977  
LOCUS BC057977 1916 bp mRNA linear ROD 08-OCT-2003  
DEFINITION Mus musculus N-acetylneuraminic acid synthase (sialic acid synthase), mRNA (cDNA clone MGC:67938 IMAGE:5052339), complete cds.  
ACCESSION BC057977  
VERSION BC057977.1 GI:37046731  
KEYWORDS MGC.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 1916)  
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toehiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, N.,

Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Greenchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Shreen, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
22388257  
12477932  
2 (bases 1 to 1916)  
Strausberg, R.  
Direct Submission  
Submitted (08-SEP-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@hgri.nih.gov](mailto:nisc_mgc@hgri.nih.gov)  
Akter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaepi, R., Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 123 Row: f Column: 14  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 16716466.  
Location/Qualifiers  
1. .1916  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="MGC:67938 IMAGE:5052339"  
/tissue\_type="Liver, normal. 5 month old male mouse."  
/clone\_lib="NCI CGAP L19"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
1. .1916  
/gene="Nans"  
/db\_xref="LocusID:94181"  
/db\_xref="MGI:2149820"  
5. .1084  
/codon\_start=1  
/product="N-acetylneuraminic acid synthase (sialic acid synthase)"  
/protein\_id="AAH57977.1"  
/db\_xref="GI:37046732"  
/translation="MPLELEICPGRWVGKHKPCFIIAEIQNHQGDIDVAKRMRTAK ECACAKFQKSELEFKFNKALERPYSKHSWGTYGEKHLRFSDQYKELQSYA QDITFTAGSGMDMAVEFLHNLNVPFFKVGSGDTNNPFLKTKAGKGRPMVTSQMG SMDTKQVQYIVKPLNPNFCLOCTSAVPLQPEDANLRVISEYQKLPDIPICYSQHE TGIASVAALGAKVLERHTLDTXWKSDBSASLEPGELAEIVRSVLEVRALGSP VLVTIETDDTVMEESVESHSKKIKA"

FEATURES  
source

gene

CDS

misc_feature		26..1057			
/note="SpsE; Region: Sialic acid synthase [Cell envelope biogenesis, outer membrane]"					
/db_xref="CDD:COG2089"					
ORIGIN					
Query Match		81.8%; Score 883.2; DB 10; Length 1916;			
Best Local Similarity		88.6%; Pred. No. 1.2e-207;			
Matches 957; Conservative		0; Mismatches 123; Indels 0; Gaps 0;			
ORIGIN					
Qy	1	ATGCCGCTGGAGCTGGAGCTGTGTCCCGGCGCTGGTGGGGGCAACACCCGCTGCTTC	60		
Db	5	ATGCCGCTGGAACTGGAGCTGTGTCCCGGCGCTGGTGGTGGAAAGCAACCCGCTGCTTC	64		
Qy	61	ATCATTTGGAGATCGGCAGACACCAACGCGGACCTGGAGCTAGCCAGCGCATGATC	120		
Db	65	ATCATTCGGGAGATCGGCAGACACCAACGCGGACATAGATGTGGCCAAAGCGCATGATC	124		
Qy	121	CGCATGGCCAGAGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAAATTC	180		
Db	125	CGCACTGCCAAGAGTGTGGGGCGGACTGGCTAAGTTTCAGAAGAGCGAGTTGGAGTTC	184		
Qy	181	AAGTTTAACTCGAAAGCCTTGGAGAGGCATACACTCGAAGCATTCCTGGGGGGAAGACG	240		
Db	185	AAGTTTAACTCGAAGGCTTGGAGAGGCATATATCTTCGAAGCATTCATGGGGGGAAGACG	244		
Qy	241	TACGGGAGCACAAGACATCTGGAGTTTCAGCCATGACCATGACAGGAGCTGCAGAGG	300		
Db	245	TATGGGAGCACACGGCGATCTGGAATTCAGCCAGCAGCATACAAAGGAGCTGCAGAGC	304		
Qy	301	TACGCCAGAGGTTGGAGTCTTCTTCACCTGCTGCGATGATGAGATGGCAGTTGAA	360		
Db	305	TATCGCAGGAGATCGGCATCTTCTTCACCTGCTGCGATGATGAGATGGCAGTTGAG	364		
Qy	361	TTCTGCTGATGAATGATGTTCCATTTTCAAAGTTGATCTGGAGACACTAATAATTTT	420		
Db	365	TTTCTGACAGACTGAATGTTCCCTTTTCAAAGTTGATCTGGAGACACTAATAATTTT	424		
Qy	421	CCTTATCTGGAAGACAGACAGCAAAAGGTGCGCCCAATGGTGATCTCCAGTGGGATGCG	480		
Db	425	CCCTACCTGGAAGACAGACAGCAAAAGGTGCTCTTATGTGATCTCCAGCGGATGCG	484		
Qy	481	TCAATGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCACTTCTGC	540		
Db	485	TCAATGACACCATGAAGCAAGTCTATCAGATCGTGAAGCCGCTGATCCCACTTCTGC	544		
Qy	541	TTCTTGAGTGTACAGCGCATACCGCTCCAGCTGAGGACGTCAACCTCGGGTCAATC	600		
Db	545	TTCTTCAATGCACAGCGGTACCCACTACAGCCGAGGATGCCAACCTCGCGGTATC	604		
Qy	601	TCGGAATATCAGAAGCTCTTCTCTGACATTCCTCAGGATTTCTGGGCATGAACAGGC	660		
Db	605	TCGGAATACAGAAGCTCTTCTCCGACATTCCTCAGGATTTCTGGGCATGAACAGGC	664		
Qy	661	ATAGCGATATCTGTGGCGCAGTGGCTCTGGGGGCCAAGTGTGTGAACTGCAATAACT	720		
Db	665	ATCGCCATATCTGTGGCGCGCTGGCTCTGGGGGCCAAGTGTGTGAACTGCAATAACG	724		
Qy	721	TTGGACAAGACCTGGGAAGGGAGTGAACACTCGGCTCGCTGGAGCTGGAGAACTGGCC	780		
Db	725	TTGGACAAGACCTGGGAAGGGAGTGAACACTCGCTCGCTGGAGCTGGGAGCTGGCA	784		
Qy	781	GAGCTGGTGGCTGAGTGGCTGTGTGGAGGTGCGCTGGAGCTGGGAGCTGGGAGCTG	840		
Db	785	GAGCTGGTGGCTGAGTGGCTGTGTGGAGGTGCGCTGGAGCTGGGAGCTGGGAGCTG	844		
Qy	841	CTGCCCTGTGAGATGGCTGCAATGAGAGTGGGCAAGTCTGTGTGGGCCCAAGTGAAG	900		
Db	845	CTGCCCTGTGAGATGGCTGCAATGAGAGTGGGCAAGTCTGTGTGGGCCCAAGTGAAG	904		
Qy	901	ATTTCGGAAGGACCACTTCTTAACATGGACATCTCACCTGGAAGTGGTGGAGCCCAA	960		
Db	905	ATCCAGCAGGACCAACCTGACCTGGACATGCTCACTGTGAAGTGGGGAGGCCCAA	964		

Qy	961	GCCTATCTCTCTGAAGACATCTTTAATCTAGTGGGCAAGAGTCTCTGCTACTGTGAA	1020			
Db	965	GGCTATCTCTCTGAAGACATCTTTCAACCTAGCGGGCAAAAGGTGCTGCTACTATCGAA	1024			
Qy	1021	GAGGATGACACATCATGGAAGAATTGGTAGTAATCATGGCAAAAAAATCAAGTCTTAA	1080			
Db	1025	GAAGATGACACGCTCATGGAAGAATCCGTGGAAAGTCAAGCAAGAAAATCAAGGCTTAA	1084			

RESULT 11					
AB041263	1174 bp	mus	linear	15-JUL-2000	
LOCUS	Mus musculus mRNA for N-acetylneuraminic acid 9-phosphate synthetase, complete cds.				
ACCESSION	AB041263				
VERSION	AB041263.1 GI:8978233				
KEYWORDS	Mus musculus (house mouse)				
SOURCE	Mus musculus				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Nakata, D., Close, B. E., Colley, K. J., Matsuda, T. and Kitajima, K. Molecular cloning and expression of the mouse N-acetylneuraminic acid 9-phosphate synthase which does not have deaminneuraminic acid (KDN) 9-phosphate synthase activity Biochem. Biophys. Res. Commun. 273 (2), 642-648 (2000)				
AUTHORS	Nakata, D., Close, B. E., Colley, K. J., Matsuda, T. and Kitajima, K.				
TITLE	Molecular cloning and expression of the mouse N-acetylneuraminic acid 9-phosphate synthase which does not have deaminneuraminic acid (KDN) 9-phosphate synthase activity Biochem. Biophys. Res. Commun. 273 (2), 642-648 (2000)				
JOURNAL	20334323				
MEDLINE	10873658				
PUBMED	2 (bases 1 to 1174)				
REFERENCE	Kitajima, K. and Nakata, D.				
AUTHORS	Direct Submission				
TITLE	Submitted (04-APR-2000) Ken Kitajima, Graduate school of bioagricultural sciences, Nagoya University, Department of applied Molecular biosciences; Chikusa-ku Huro-u-chou, Nagoya, Aichi 464-8601, Japan (E-mail: kitajima@agr.nagoya-u.ac.jp, Tel:+81-52-789-4130, Fax:+81-52-789-4128)				
JOURNAL	Location/Qualifiers				
MEDLINE	1..1174				
PUBMED	/organism="Mus musculus"				
REFERENCE	/mol_type="mRNA"				
AUTHORS	/db_xref="taxon:10090"				
TITLE	/cell_line="COMMALD"				
JOURNAL	71..1150				
FEATURES	/codon_start=1				
source	/product="N-acetylneuraminic acid 9-phosphate synthetase"				
	/protein_id="BAA98131.1"				
	/db_xref="GI:8978233"				
CDS	/translation="MPLELELCPRWVGKHPFIIAIEIGNHGGIDIVAKEMIRTAKECGADCAKFKSELEPFKRNKALERTPYTSKXSGKTYGEHKHLEFHDQYKEIQSYA QEIGIFFTASGMDMAVEFLHNLNVPFVKVSGDTNNFVLEKTKAKRPMVIGSGMO SMDTMKQVQIVKPLNPNFCFLOCTSAVPLQPEDANLRVISEYOKLPDIPIGYSGHE TGAISVAVALGAKVLERHTLTDKTKWGDHSLSPGLAEALVRSVLVERALGSP TKOLLPCEMACNEKLGKSVAKVKIPAGTILTLTLTVKVGEPKYPEDIFNLGAKK VLVTIEDDITVMEESVESHSKIKKA"				

Query Match	81.6%;	Score	881.6;	DB	10;	Length	1174;
Best Local Similarity	88.5%;	Pred. No.	2.9e-207;				
Matches	956;	Conservative	0;	Mismatches	124;	Indels	0;
Gaps	0;						

Qy	1	ATGCCGCTGGAGCTGGAGCTGTGTCCCGGCGCTGGTGGGGGCAACACCCGCTGCTTC	60			
Db	71	ATGCCGCTGGAACTGGAGCTGTGTCCCGGCGCTGGTGGTGGAAAGCAACCCGCTGCTTC	130			
Qy	61	ATCATTTGGAGATCGGCAGACACCAACGCGGCGACCTGGAGCTAGCCAGCGCATGATC	120			
Db	131	ATCATTCGGGAGATCGGCAGACACCAACGCGGACATAGATGTGGCCAAAGCGCATGATC	190			
Qy	121	CGCATGGCCAGAGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGCTAGAAATTC	180			
Db	191	CGCACTGCCAAGAGTGTGGGGCGGACTGCTGCTAAGTTTCAGAAGAGCGAGTTGGAGTTC	250			

```
QY 181 AAGTTTAAATCGAAGAGCTTGGAGAGGCCATACACTCGAAGCATTCCTCGGGGGAAGACG 240
Db 251 AAGTTTAAACCGAAGAGCTTGGAGAGCCATATATCTCGAAGCATTCATGGGGGAAGACG 310
QY 241 TAACGGGAGCACAACGACATCTGAGTTACGCCATGACCACTAGCAGGAGCTGCAGAGG 300
Db 311 TATGGGAGCACAACGCGCATCTGGAATTCAGCCAGCCAGGATCAAGAGAGCTGCAGAGC 370
QY 301 TACGCCGAGGAGTTGGGATCTTCTCACTGCTCTGGCATGGATGAGATGCGACGTTGAA 360
Db 371 TATGGCAGGAGATCGGCATCTTCTCACTGCTCTGGCATGGAGAGATGCGACGTTGAG 430
QY 361 TTCTCGCATGAATGAATGTTCCATTTTCAAAGTTGGATCTGGAGACACATTAATTTT 420
Db 431 TTCTCGCAGCACTGAATGTTCCCTTTTCAAAGTTGGATCTGGGGGACACTAACCAATTTT 490
QY 421 CCTTATCTGGAAGACAGCAGCAAAAGTGCCTCAATGGTGATCTCCAGTGGGATGCAG 480
Db 491 CCCTACCTGGAAAGACAGCCAGAAAGTGTCTTATGGTGATCTCCAGCGGATGCAG 550
QY 481 TCAATGGACACCATGAAGCAAGTTTATCAGATCTGTAAGCCCTCAACCCCAACTTCTGC 540
Db 551 TCAATGGACACCATGAAGCAAGTCTATCAGATCTGTAAGCCCTCAATCCCAACTTCTGC 610
QY 541 TTCTTGCAGTGATACAGCGCATACCGCTCCAGCTGAGGAGCTCAACCTGCGGCTCATC 600
Db 611 TTCTTCCAAATGACAGCGGTACCACTTACAGCCGAGGATGCCAACTGCGGCTCATC 670
QY 601 TCGGAATATCAGAAGCTCTTCTCGACATTTCCATAGGTAATCTCGGGCATGAAACAGGC 660
Db 671 TCGGAATATCAGAAGCTCTTCTCGACATTTCCATAGGTAATCTCGGGCATGAAACAGGC 730
QY 661 ATAGCGATATCTGTGGCGCAGTGTCTGTGGGGCCAAAGTGTGTGAAGCTCACATACT 720
Db 731 ATGCGCATATCTGTGGCGCGCTGTGTGGGGCCAAAGTGTGTGAAGCTCACATAAG 790
QY 721 TTGGACAGACCTGGAAGGGAGTGCACCTCGGCTCGCTCGGAGCTTGGAGCTTGGAGCTG 780
Db 791 TTGGACAGACCTGGAAGGGAGTGCACCTCGGCTCGCTCGGAGCTTGGAGCTTGGAGCTG 850
QY 781 GAGCTGTGTGGTCAAGTGTCTGTGGAGCGTGTCTGTGGGCTGCCCAACCAAGCAGCTG 840
Db 851 GAGCTGTGTGGTCAAGTGTCTGTGGAGCGTGTCTGTGGGCTGCCCAACCAAGCAGCTG 910
QY 841 CTGCGCTGTGAGATGGCTGTCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCCAAGTGAAA 900
Db 911 CTGCGCTGTGAGATGGCTGTCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCCAAGTGAAA 970
QY 901 ATTCCGGAGGACCATTTCTACAAATGGACATGCTCACCGTGAAGTGGGTGAGCCCAA 960
Db 971 ATCCAGAGGACCATTCACCTGACCTTGGACAGCTCACTGTGAAGTGGGGAGCCCAA 1030
QY 961 GCCTATCTCTCTGAAGACATCTTTAATCTAGTGGGCAAGAGTCTCTGTGCTCACTGTGAA 1020
Db 1031 GCCTATCTCTCTGAAGACATCTTTAATCTAGTGGGCAAGAGTCTCTGTGCTCACTGTGAA 1090
QY 1021 GAGGATGACACCATCATGGAAGAAATGGTAGAATATCATGGCAAAAATCAAGTCTTAA 1080
Db 1091 GAGGATGACACCGTCAATGGAAGAAATCCGTGGAAAGTCAAGCAAGAAAATCAAGGCTTAA 1150

RESULT 12
AK091184
LOCUS
DEFINITION
Homo sapiens cDNA FLJ33865 fis, clone CTONG2006515, highly similar
to N-acetylneuraminic acid phosphate synthase.
ACCESSION
AK091184
VERSION
AK091184.1 GI:21749493
KEYWORDS
oligo capping; fis (full insert sequence).
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
REFERENCE
AUTHORS
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 Nishi, T., Ota, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 3026)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kasuga-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
source
Location/Qualifiers
1..3026
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CTONG2006515"
/tissue_type="tongue, tumor tissue"
/clone_lib="CTONG2"
/note="cloning vector: pME18SFL3"
ORIGIN
Query Match 67.8%; Score 732.4; DB 9; Length 3026;
Best Local Similarity 99.9%; Pred. No. 2.7e-170;
Matches 733; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 347 AGATGGCAGTTGAATTCCTGCATGAATCTGAATGTTCCATTTTCAAAGTTGATCTGGAG 406
Db 2273 AGATGGCAGTTGAATTCCTGCATGAATCTGAATGTTCCATTTTCAAAGTTGATCTGGAG 2332
QY 407 ACACATAAATTTTCTCTTATCTGGAAGACAGCCAAAAGGTCGCCAATGGTGATCT 466
Db 2333 ACACATAAATTTTCTCTTATCTGGAAGACAGCCAAAAGGTCGCCAATGGTGATCT 2392
QY 467 CCAGTGGGATGCAGTCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCA 526
Db 2393 CCAGTGGGATGCAGTCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCA 2452
QY 527 ACCCAAATCTTGTCTTTCAGTGTACAGCGATACCCGCTCCAGCTCGAGGAGCTCA 586
Db 2453 ACCCAAATCTTGTCTTTCAGTGTACAGCGATACCCGCTCCAGCTCGAGGAGCTCA 2512
QY 587 ACCTGGGGTCACTCTCGGAATATCAGAGCTCTTTCTGCACATTCCTCAGGCTATTCG 646
Db 2513 ACCTGGGGTCACTCTCGGAATATCAGAGCTCTTTCTGCACATTCCTCAGGCTATTCG 2572
QY 647 GGCATGAACACGCGATAGCGATATCTGTGGCCGAGTGGCTCTGGGGGCAAGGTTGG 706
Db 2573 GGCATGAACACGCGATAGCGATATCTGTGGCCGAGTGGCTCTGGGGGCAAGGTTGG 2632
QY 707 AACGTCACTAATCTTGGACAAGACCTGGAAGGGAGTGACCACTCGGCTCGCTGGAGC 766
Db 2633 AACGTCACTAATCTTGGACAAGACCTGGAAGGGAGTGACCACTCGGCTCGCTGGAGC 2692
QY 767 CTGGAGAACTGCGCGAGCTGGTGGCTCAGTGCCTCTTCTGGAGCGCTGCCCTGGGCTCCC 826
Db 2693 CTGGAGAACTGCGCGAGCTGGTGGCTCAGTGCCTCTTGTGGAGCGCTGCCCTGGGCTCCC 2752
QY 827 CAACCAAGCAGCTGTGCGCTCTGTGAGATGGCTGCAATGAGAAGCTGGGCAAGCTGTGG 886
```

```
Db 2753 CAACCAAGCAGCTGCTGCCCTGTGAGTGGCCCTGCAATGAGAGCTGGGCAAGTCTGTGG 2812
Qy 887 TGGCCAAAGTGAATTCGGGAAGGACCAATCTTAACAATGACATGCTCAACCGTGAAGG 946
Db 2813 TGGCCAAAGTGAATTCGGGAAGGACCAATCTTAACAATGACATGCTCAACCGTGAAGG 2872
Qy 947 TGGGTGAGCCCAAGCTATCTCTCTGAAGACATCTTTAATCTAGTGGGCAAGAGGTCC 1006
Db 2873 TGGGTGAGCCCAAGCTATCTCTCTGAAGACATCTTTAATCTAGTGGGCAAGAGGTCC 2932
Qy 1007 TGGTCACTGTTGAAGAGGATGACACCATCATGGAAGAAATGGTAGATAATCATGSCAAA 1066
Db 2933 TGGTCACTGTTGAAGAGGATGACACCATCATGGAAGAAATGGTAGATAATCATGSCAAA 2992
Qy 1067 AAATCAAGTCITAA 1080
Db 2993 AAATCAAGTCITAA 3006

RESULT 13
AX868777
LOCUS AX868777 813 bp DNA linear PAT 17-DEC-2003
DEFINITION Sequence 3682 from Patent EP1074617.
ACCESSION AX868777
VERSION AX868777.1 GI:40023640
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
Primers for synthesizing full-length cDNA and their use
Patent: EP 1074617-A 3682 07-FEB-2001;
Research Association for Biotechnology (JP)
FEATURES
source
1. 813
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 65.8%; Score 710.2; DB 6; Length 813;
Best Local Similarity 98.7%; Pred. No. 7.8e-165;
Matches 734; Conservative 0; Mismatches 8; Indels 2; Gaps 2;
Qy 1 ATGCCGCTGGAGCTGGAGCTGTGTCCTCCGGGGCGCTGGGTGGGGGCAACACCCGCTGCTTC 60
Db 71 ATGCCGCTGGAGCTGGAGCTGTGTCCTCCGGGGCGCTGGGTGGGGGCAACACCCGCTGCTTC 130
Qy 61 ATCATTCGCGAGATCGGCAGACACACAGGGCGACCTGGACGTAGCCCAAGCGCATGATC 120
Db 131 ATCATTCGCGAGATCGGCAGACACACAGGGCGACCTGGACGTAGCCCAAGCGCATGATC 190
Qy 121 CGCATGGCCAAAGGAGTGGGGCTGATTGTGCCAAGTTCACAGAGAGTGGAGTGAATTC 180
Db 191 CGCATGGCCAAAGGAGTGGGGCTGATTGTGCCAAGTTCACAGAGAGTGGAGTGAATTC 250
Qy 181 AAGTTTAAATCGAAAGCCTTGGAGAGGCGCATACACCTCGAAGCATTCCTCGGGGAAGACG 240
Db 251 AAGTTTAAATCGAAAGCCTTGGAGAGGCGCATACACCTCGAAGCATTCCTCGGGGAAGACG 310
Qy 241 TACGGGAGCAAAACGACATCTGGAGTTACGCAATGACCATGACAGTACAGGAGCTCAGAGG 300
Db 311 TACGGGAGCAAAACGACATCTGGAGTTACGCAATGACCATGACAGTACAGGAGCTCAGAGG 370
Qy 301 TACGCCGAGGAGGTGGGATCTTCTTCACTGCTCTGCGATGATGATGATGATGATGATGATG 360
Db 371 TACGCCGAGGAGGTGGGATCTTCTTCACTGCTCTGCGATGATGATGATGATGATGATGATG 430
Qy 361 TTCCTGCATGAACCTGAATGTTCCATTTTTCAAAGTTGGATCTGGAGAGCACTAAATTTT 420
```

```
Db 431 TTCCTGCATGAACCTGAATGTTCCATTTTCAAGTTGGATCTGGAGACACTAATAATTTT 490
Qy 421 CCTTATCTGGAAGACAGCCAAAGAGTCCCAATGGTGATCTCCAGTGGGATGCAG 480
Db 491 CCTTATCTGGAAGACAGCAGNCAAAAAGGTCCGCCAATGGTGATCTCCAGTGGGATGCAG 550
Qy 481 TCATGGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCCAATCTTGC 540
Db 551 TCAATGGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCCAATCTTGC 610
Qy 541 TTCCTGCAGTGTACAGCGCATACCCGCTCCAGCGCTGAGGAGCTCAACCTCGGGGTCATC 600
Db 611 TTCCTGCAGTGTACAGCGCATACCCGCTCCAGCGCTGAGGAGCTCAACCTCGGGGTCATC 670
Qy 601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCTGAGGCTATCTGGGCGATGAACAGGC 660
Db 671 TCGGAATATCAGAAGCTCTTTCCTGACATTCCTGAGGCTATCTGGGCGATGAACAGGC 730
Qy 661 ATAGCGATATCTGTGGCCGCGAGTGGC-TCTGGGGGCCAAGGTGTTGGAACTGCAATAC 719
Db 731 ATAGCGATATCTGTGGCCGCGAGTGGC-TTGGGGACCAAGGTGNTGGAACGTCAAT-AC 789
Qy 720 TTTGGACAAGACCTTGGAAAGGGAG 743
Db 790 TTTGGACAAGACCTTGGAAAGGGAG 813

RESULT 14
BD148839
LOCUS BD148839 813 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD148839
VERSION BD148839.1 GI:27854597
KEYWORDS JP 2002191363-A/3682.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 813)
Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 3682 09-JUL-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/3682
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
SAITO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT source 1. 813
/organism="Homo sapiens (human)".
FT source
1. 813
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 65.8%; Score 710.2; DB 6; Length 813;
Best Local Similarity 98.7%; Pred. No. 7.8e-165;
Matches 734; Conservative 0; Mismatches 8; Indels 2; Gaps 2;
Qy 1 ATGCCGCTGGAGCTGGAGCTGTGTCCTCCGGGGCGCTGGGTGGGGGCAACACCCGCTGCTTC 60
```

Db 71 ATGCGCTGGAGCTGGAGCTGTGTCCTGGGGCGCTGGGTGGGGGNCACACCCCTGGCTTC 130  
Qy 61 ATCATTTGCCGAGATCGCCAGAACACACAGGGCGACCTGGACGTAGCCAAAGCGCATGATC 120  
Db 131 ATCATTTGCCGAGATCGCCGAGAACACACAGGGCGACCTGGACGTAGCCAAAGCGCATGATC 190  
Qy 121 CGCATGGCCAGGAGTGTGGGGCTGATTTGTCCCAAGTTCGAAAGAGTGAAGTGAATTC 180  
Db 191 CGCATGGCCAGGAGTGTGGGGCTGATTTGTCCCAAGTTCGAAAGAGTGAAGTGAATTC 250  
Qy 181 AAGTTTAATCGAAAGCCCTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240  
Db 251 AAGTTTAATCGAAAGCCCTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 310  
Qy 241 TACGGGGAGCACAACAGCATCTGGAGTTCAGCCCATGACCGATACAGGGAGCTGCAGAGG 300  
Db 311 TACGGGGAGCACAACAGCATCTGGAGTTCAGCCCATGACCGATACAGGGAGCTGCAGAGG 370  
Qy 301 TACGCCGAGGAGTGGGATCTTCTTCACTGCCTCGCATGGATGAGATGCGAGTTCGAA 360  
Db 371 TACGCCGAGGAGTGGGATCTTCTTCACTGCCTCGCATGGATGAGATGCGAGTTCGAA 430  
Qy 361 TTCTGTCATGAATGAATGTTCCATTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420  
Db 431 TTCTGTCATGAATGAATGTTCCATTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 490  
Qy 421 CTCTATCTGAAAAGACAGCCAAAAGGTGCGCCCAATGGTGATCTCCAGTGGGATGACG 480  
Db 491 CTCTATCTGAAAAGACAGCCAAAAGGTGCGCCCAATGGTGATCTCCAGTGGGATGACG 550  
Qy 481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTGC 540  
Db 551 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTGC 610  
Qy 541 TTCTGTCAGTGTACAGCGCATACCGCTCCAGCTCGAGGAGTCAACCTCGGGTCAATC 600  
Db 611 TTCTGTCAGTGTACAGCGCATACCGCTCCAGCTCGAGGAGTCAACCTCGGGTCAATC 670  
Qy 601 TCGGAATATCAGAGCTCTTCTGACATTCGCATAGGATCTTGGGCGATGAACAGGC 660  
Db 671 TCGGAATATCAGAGCTCTTCTGACATTCGCATAGGATCTTGGGCGATGAACAGGC 730  
Qy 661 ATAGCGATATCTGCGCGCAGTGC-TCTGGGGCCCAAGGTGTTGGAACGTCACATAAC 719  
Db 731 ATAGCGATATCTGCGCGCAGTGC-TTGGGGACCAAGGTGTTGGAACGTCACAT-AC 789  
Qy 720 TTTGGACAAGACCTGGAAGGGGAG 743  
Db 790 TTTGGACAAGACCTGGAANGGAG 813

RESULT 15  
AX873806/c  
LOCUS AX873806 514 bp DNA linear PAT 17-DEC-2003  
DEFINITION Sequence 8711 from Patent EP1074617.  
ACCESSION AX873806  
VERSION AX873806.1 GI:40028595  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE  
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,  
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.  
TITLE Primers for synthesising full-length cDNA and their use  
JOURNAL Patent: EP 1074617-A 8711 07-FEB-2001;  
Research Association for Biotechnology (JP)  
FEATURES  
Location/Qualifiers  
1..514  
source  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"  
ORIGIN  
Query Match 44.8%; Score 483.8; DB 6; Length 514;  
Best Local Similarity 98.2%; Pred. No. 8.6e-109;  
Matches 485; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
Qy 587 ACCTGGCGGTTCATCTCGGAATATCAGAGCTCTTTCTGAGCATTCCTGAGGATTCCTGAGGATTCCTG 646  
Db 514 ACNTGNGGGTNNNTNGGAANATCAGAAAGCTCTTTCTGACATTCCTGAGGATTCCTGAGGATTCCTG 455  
Qy 647 GGCAATGAAAACAGGCATAGCGATATCTGTGGCCGCGAGTGGCTCTGCGGGCCCAAGGTGTTGG 706  
Db 454 GGCAATGAAAACAGGCATAGCGATATCTGTGGCCGCGAGTGGCTCTGCGGGCCCAAGGTGTTGG 395  
Qy 707 AACGTCAATAACTTTGGACAAGACCTGGAAAGGGAGTGACCACTCGGCGCTCGCTGGAGC 766  
Db 394 AACGTCAATAACTTTGGACAAGACCTGGAAAGGGAGTGACCACTCGGCGCTCGCTGGAGC 335  
Qy 767 CTGGAGAACTGGCCGAGCTGGTGGCTCAGTGCCTTTGTGGAGCGTCCCTGGGCTCCC 826  
Db 334 CTGGAGAACTGGCCGAGCTGGTGGCTCAGTGCCTTTGTGGAGCGTCCCTGGGCTCCC 275  
Qy 827 CAACCAAGCAGCTGCTGCCCTGTGAGATGGCTTGCATGAGAAAGCTGGGCAAGTCTGTGG 886  
Db 274 CAACCAAGCAGCTGCTGCCCTGTGAGATGGCTTGCATGAGAAAGCTGGGCAAGTCTGTGG 215  
Qy 887 TGGCCAAAGTGAAAATTCGGGAAGCCACCATCTTAACAATGGACATGCTCACCCGTGAAGG 946  
Db 214 TGGCCAAAGTGAAAATTCGGGAAGCCACCATCTTAACAATGGACATGCTCACCCGTGAAGG 155  
Qy 947 TGGGTGAGCCCAAGCCCTATCCTCTGAAGACATCTTTAATCTAGTGGGCAAGAGGTCC 1006  
Db 154 TGGGTGAGCCCAAGCCCTATCCTCTGAAGACATCTTTAATCTAGTGGGCAAGAGGTCC 95  
Qy 1007 TGGTCACTGTTGAAGAGGATGACCACTATGGAAGAAATGGTAGATAATCATGCGCAAAA 1066  
Db 94 TGGTCACTGTTGAAGAGGATGACCACTATGGAAGAAATGGTAGATAATCATGCGCAAAA 35  
Qy 1067 AAATCAAGTCTTAA 1080  
Db 34 AAATCAAGTCTTAA 21

Search completed: September 16, 2004, 04:04:22  
Job time : 6858 sec

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2004, 21:22:11 ; Search time 668 Seconds  
(without alignments)  
6868.344 Million cell updates/sec

Title: US-10-759-277-3  
Perfect score: 1080  
Sequence: 1 atgcgcgtggagctggagct.....gcaaaaaaatcaagctcttaa 1080

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_29Jan04: \*  
1: geneseqn1980s: \*  
2: geneseqn1990s: \*  
3: geneseqn2000s: \*  
4: geneseqn2001as: \*  
5: geneseqn2001bs: \*  
6: geneseqn2002s: \*  
7: geneseqn2003as: \*  
8: geneseqn2003bs: \*  
9: geneseqn2003cs: \*  
10: geneseqn2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1080	100.0	1080	3 AAA37763	Aaa37763 Human gly
2	1080	100.0	1080	4 AAH28458	Aah28458 Nucleotid
3	1080	100.0	1080	7 AAL53993	Aal53993 DNA encod
4	1080	100.0	1146	3 AAA50569	Aaa50569 Human sia
5	1078	99.8	1268	6 ABL90630	AbL90630 Human pol
6	1076.8	99.7	1230	4 AAI59142	Aai59142 Human pol
7	1076.8	99.7	1230	8 ADB49125	Adb49125 Novel hum
8	1075.2	99.6	1170	4 AAH14557	Aah14557 Human cDN
9	898.2	83.2	1243	4 AAI60928	Aai60928 Human pol
10	732	67.8	1238	4 AAS33141	Aas33141 DNA encod
11	710.2	65.8	813	4 AAH06847	Aah06847 Human cDN
12	483.8	44.8	514	4 AAH11876	Aah11876 Human cDN
13	426.4	39.5	1417	5 AAS91837	Aas91837 DNA encod
14	418	38.7	464	6 ABN95016	Abn95016 Gene #151
15	386.8	35.8	553	6 ABK94731	Abk94731 Neurodege
16	371.8	34.4	486	8 ACH18008	Ach18008 Human adu
17	336.8	31.2	442	8 ACH41408	Ach41408 Human foe
18	327.8	30.4	417	8 ACH38131	Ach38131 Human end
19	315.8	29.2	480	8 ACH32141	Ach32141 Human end
20	280.6	26.0	404	5 AAH42798	Aah42798 Nucleotid
21	280.6	24.1	1028	5 AAS81438	Aas81438 DNA encod
22	256.2	23.7	410	4 AAH35557	Aah35557 Human col
23	198.2	18.4	1017	4 ABL29095	AbL29095 Drosophil

24	162.4	15.0	611	5 AAS91836	Aas91836 DNA encod
25	146	13.5	492	4 AAK88570	Aak88570 Human dig
26	119.8	11.1	3123	4 ABL29094	AbL29094 Drosophil
27	119.8	11.1	3375	4 ABL29196	AbL29196 Drosophil
28	112.4	10.4	26281	3 AAZ60929	Aaz60929 Nucleotid
29	101.6	9.4	17276	7 ACA64723	Aca64723 Streptoco
30	100	9.3	1023	6 ABN67441	Abn67441 Streptoco
31	100	9.3	110000	6 ABN71527_11	Continuation (12 o
32	90	8.3	3777	4 ABL29200	AbL29200 Drosophil
33	70.4	6.5	110000	2 AAV21209_04	Continuation (5 of
34	63	5.8	1059	3 AAA50570	Aaa50570 Escherich
35	63	5.8	1059	4 AAH28459	Aah28459 Nucleotid
36	63	5.8	1059	7 AAL53994	Aal53994 DNA encod
37	62.8	5.8	999	4 AAP91405	Aap91405 N. mening
38	62.8	5.8	1000	6 ABK37785	Abk37785 DNA sequ
39	62.8	5.8	1050	6 ABA99765	AbA99765 N. mening
40	62.8	5.8	65632	3 AAA81502	Aaa81502 N. mening
41	62.8	5.8	110000	3 AAB81490_00	Aab81490 N. mening
42	62.8	5.8	349980	3 AAF21544	Aaf21544 Neisseria
43	62	5.7	133	3 AAC98691	Aac98691 Human col
44	60	5.6	60	6 ABN36327	Abn36327 Human spl
45	59.6	5.5	984	6 ABV72354	Abv72354 Nucleotid

## ALIGNMENTS

RESULT 1					
AAA37763	AAA37763	standard; DNA; 1080 BP.			
XX	XX	AAA37763;			
XX	XX	04-DEC-2000 (first entry)			
DT	DT	Human glycosylation enzyme clone HASAA37 coding sequence.			
XX	XX				
DE	DE	Human; glycosylation enzyme; glycolysis; myoglobinuria; tumour marker; immunotherapy; cosmetic surgery; metabolism; immune system disorder; haematopoietic cell deficiency; blood coagulation disorder; asthma; afibrinogenemia; blood platelet disorder; thrombocytopaenia; neoplasia; autoimmune disorder; Addison's disease; multiple sclerosis; purpura; allergic encephalomyelitis; allergic reaction; organ rejection; graft-versus-host disease; inflammation; hyperproliferative disorder; sarcoidosis; infection; gene therapy; CMP sialic acid synthetase; ds.			
KW	KW	Homo sapiens.			
XX	XX				
XX	XX	Key	Location/Qualifiers		
FT	FT	CDS	1..1080		
FT	FT		/*tag= a		
XX	XX		/product= "Sialic_acid_synthetase"		
XX	XX	WO2000052136-A2.			
XX	XX	08-SEP-2000.			
XX	XX	01-MAR-2000; 2000WO-US005325.			
XX	XX	02-MAR-1999; 99US-0122409P.			
XX	XX	(HUMA-) HUMAN GENOME SCI INC.			
XX	XX	Coleman TA;			
XX	XX	WPI: 2000-572179/53.			
XX	XX	P-PSDB; AAY90352.			
XX	XX	New human glycosylation enzymes cytidine 5'-monophosphate sialic acid synthetase, sialic acid synthetase and aldolase and nucleic acids encoding the proteins for treating e.g., immune system disorders, microbial diseases.			







XX Betenbaugh MJ, Lawrence S, Lee YC, Coleman TA;  
 XX WPI; 2003-102519/09.  
 XX P-PSDB; AAO26546.  
 XX Manipulating glycoprotein production in insect cell, involves enhancing  
 PT expression of enzymes involved in carbohydrate processing pathway such as  
 PT N-acetylglucosamine-2 epimerase or sialic acid synthetase.  
 XX Disclosure; Fig 31; 88pp; English.  
 XX The invention relates to a novel method for manipulating glycoprotein  
 CC production in an insect cell comprising enhancing expression of an  
 CC enzyme, such as N-acetylglucosamine-2 (GlcNAc-2) epimerase, one  
 CC catalyzing conversion of UDP-GlcNAc to mannose (Man)NAc, sialic acid  
 CC synthetase, aldolase, cytidine monophosphate-sialic acid (CMP-SA)  
 CC synthetase or CMP-SA transporter, where the expression of each enzyme is  
 CC enhanced to above endogenous levels. The novel method is useful for  
 CC manipulating glycoprotein production in an insect cell. Further methods  
 CC of the invention are useful for producing sialylated glycoprotein. The  
 CC sialylated glycoprotein produced by the above mentioned methods are  
 CC useful as pharmaceutical compositions, vaccines, diagnostics and  
 CC therapeutics. This polynucleotide sequence represents the DNA encoding a  
 CC human SA synthetase protein of the invention  
 XX  
 SQ Sequence 1080 BP; 274 A; 261 C; 314 G; 231 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 1080; DB 7; Length 1080;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGCGCTGGAGCTGGAGCTGTGTCCTGGGCGCTGGTGGGGGCAACACCCGCTTC 60  
 Db 1 ATGCGCTGGAGCTGGAGCTGTGTCCTGGGCGCTGGTGGGGGCAACACCCGCTTC 60  
 QY 61 ATCATTTGCGGAGATCGGCAGAACACACAGGCGACCTGGAGCTAGCCAGCGCATGATC 120  
 Db 61 ATCATTTGCGGAGATCGGCAGAACACACAGGCGACCTGGAGCTAGCCAGCGCATGATC 120  
 QY 121 CGCATGGCCAGGAGTGTGGGCTGATTTGTCGCAAGTTCCAGAGAGTGTAGAAATTC 180  
 Db 121 CGCATGGCCAGGAGTGTGGGCTGATTTGTCGCAAGTTCCAGAGAGTGTAGAAATTC 180  
 QY 181 AAGTTTAATCGAAAGCCCTTGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAGAGC 240  
 Db 181 AAGTTTAATCGAAAGCCCTTGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAGAGC 240  
 QY 241 TACGGGGAGCACAAGCATCTGGAGTTTCAGCCATGACAGGAGCTGACAGAGG 300  
 Db 241 TACGGGGAGCACAAGCATCTGGAGTTTCAGCCATGACAGGAGCTGACAGAGG 300  
 QY 301 TACGGCGAGGAGTTGGGATCTTTCTCACTGCTCTGGCATGGATGAGATGGCAGTTGAA 360  
 Db 301 TACGGCGAGGAGTTGGGATCTTTCTCACTGCTCTGGCATGGATGAGATGGCAGTTGAA 360  
 QY 361 TTCTGCGATGAATGAATGTTCCATTTTCAAGTTGGATCTGGAGACACTAATAATTTT 420  
 Db 361 TTCTGCGATGAATGAATGTTCCATTTTCAAGTTGGATCTGGAGACACTAATAATTTT 420  
 QY 421 CTTTATCTGGAAGAAGACCCCAAAAGGTGCGCCAAATGGTGTATCTCCAGTGGATGACG 480  
 Db 421 CTTTATCTGGAAGAAGACCCCAAAAGGTGCGCCAAATGGTGTATCTCCAGTGGATGACG 480  
 QY 481 TCAATGGACACCATGAAGCAAGTTTATCAGATCTGTGAAGCCCTCAACCCCACTTCTGC 540  
 Db 481 TCAATGGACACCATGAAGCAAGTTTATCAGATCTGTGAAGCCCTCAACCCCACTTCTGC 540  
 QY 541 TTCTTGCAGTGTACAGCGGATACCCGCTCCAGCTGAGGAGCGTCAACCTGGGGTCATC 600  
 Db 541 TTCTTGCAGTGTACAGCGGATACCCGCTCCAGCTGAGGAGCGTCAACCTGGGGTCATC 600  
 QY 601 TCGGAATATCAGAAGCTCTTTCTCTGACATTCCTCCATAGGGTATTCTGGGCATGAACAGCG 660

Db 601 TCGGAATATCAGAAGCTCTTTCTCTGACATTCCTCCATAGGGTATTCTGGGCATGAACAGCG 660  
 QY 661 ATAGCGATATCTGTGGCGCAGTGGCTCTGGGGGCAAGGTTGGTGGAAACGTCACATAACT 720  
 Db 661 ATAGCGATATCTGTGGCGCAGTGGCTCTGGGGGCAAGGTTGGTGGAAACGTCACATAACT 720  
 QY 721 TTGGAACAAGCTCGAAGGGAGTGACACACTCGGCTCGCTGGAGCCCTGGAGACTGGCC 780  
 Db 721 TTGGAACAAGCTCGAAGGGAGTGACACACTCGGCTCGCTGGAGCCCTGGAGACTGGCC 780  
 QY 781 GAGCTGGTGGCGTCAAGTGGCTCTTTGTGAGAGCTGCTTGGGCTCCCAACCAAGCAGCTG 840  
 Db 781 GAGCTGGTGGCGTCAAGTGGCTCTTTGTGAGAGCTGCTTGGGCTCCCAACCAAGCAGCTG 840  
 QY 841 CTGCCCTGTGAGATGGCTCGCAATGAGAGCTGGCAAGTCTGTGGTGGCCAAAGTGAAA 900  
 Db 841 CTGCCCTGTGAGATGGCTCGCAATGAGAGCTGGCAAGTCTGTGGTGGCCAAAGTGAAA 900  
 QY 901 ATTCGGGAAGCACCATTCTAACCAATGACATGCTCACCGTGAAGTGGGTGAGCCAAA 960  
 Db 901 ATTCGGGAAGCACCATTCTAACCAATGACATGCTCACCGTGAAGTGGGTGAGCCAAA 960  
 QY 961 GCCTATCTCTCTGAAGACATCTTTAATCTAGTGGCAAGAGGTCCTGTCTCACTGTGAA 1020  
 Db 961 GCCTATCTCTCTGAAGACATCTTTAATCTAGTGGCAAGAGGTCCTGTCTCACTGTGAA 1020  
 QY 1021 GAGGATGACACCATCATGGAAGAAATTGGTAGATATCATGGCAAAAAAATCAAGTCTTAA 1080  
 Db 1021 GAGGATGACACCATCATGGAAGAAATTGGTAGATATCATGGCAAAAAAATCAAGTCTTAA 1080  
 RESULT 4  
 AAA50569  
 ID AAA50569 standard; cDNA; 1146 BP.  
 XX AC AAA50569;  
 XX 19-DEC-2000 (first entry)  
 XX Human sialic acid synthetase cDNA.  
 XX Sialic acid synthetase; human; sas gene; sialylation; glycoprotein;  
 KW plasminogen; transferrin; thyrotropin; Na+,K+-ATPase; ss.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 FT 46..1125  
 CDS /\*tag= a  
 XX WO200052135-A2.  
 XX 08-SEP-2000.  
 XX 01-MAR-2000; 2000WO-US005313.  
 XX 02-MAR-1999; 99US-0122582P.  
 PR 08-DEC-1999; 99US-0169624P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA (UYJO ) UNIV JOHNS HOPKINS.  
 PA (UYWY-) UNIV WYOMING.  
 XX Betenbaugh MJ, Lawrence S, Lee YC, Jarvis D, Coleman TA;  
 XX WPI; 2000-572178/53.  
 DR P-PSDB; AAY96101.  
 XX Recombinant production of sialylated glycoproteins using cells in which  
 PT the expression of enzymes, e.g. sialic acid synthetase, involved in the  
 PT sialylation reaction has been altered.  
 XX

Disclosure; Page 103-105; 144bp; English.

The present sequence is that of human cDNA encoding sialic acid synthetase (see AY96101), an enzyme that condenses ManNAc-6-P or Man-6-P with FBP to form Neu5Ac and KDN phosphates, respectively. This novel gene, termed SAS, was identified on the basis of homology with the *Escherichia coli* neuB gene using a bioinformatics based approach. Northern blots indicated ubiquitous transcription of the human synthetase gene in a selection of tissues. The invention provides methods and recombinantly engineered cells for producing glycoproteins having sialylated oligosaccharides. The methods involve altering the expression of enzymes involved in carbohydrate processing. A claimed cell producing sialylated glycoprotein at above endogenous levels expresses at least 1 (preferably human) enzyme selected from GlcNAc-2 epimerase, an enzyme catalyzing the conversion of UDP-GlcNAc to ManNAc, sialic acid synthetase, aldolase, CMP-sialic acid synthetase and CMP-sialic acid transporter at above endogenous levels. Endogenous N-acetylglucosaminidase activity may be suppressed. A claimed method for manipulating glycoprotein in an insect cell comprises enhancing the expression of 1 of the above enzymes, and a claimed method for producing sialylated glycoproteins involves expressing a heterologous protein (especially plasminogen, transferrin, Na<sup>+</sup>/K<sup>+</sup>-ATPase or thyrotropin) in the insect cell. Yeast, insect, fungal, plant and bacterial host cells can be engineered to produce new forms of sialylated glycoproteins, higher concentrations of sialylated glycoproteins and/or elevated concentrations of donor substrates (e.g. nucleotide sugars) required for sialylation

Sequence 1146 BP; 290 A; 280 C; 332 G; 244 T; 0 U; 0 Other;

Query Match 100.0%; Score 1080; DB 3; Length 1146;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGCGCTGGAGCTGGAGCTGTGTCCTGGGGCTGGGGTGGGGCGGCAACACCGGTGCTC 60  
 46 ATGCGCTGGAGCTGGAGCTGTGTCCTGGGGCTGGGGTGGGGCGGCAACACCGGTGCTC 105  
 61 ATCATTTCCGAGATCGGCCAGAACACACGAGCGGACCTGGAGCTAGCCAAAGCGATGATC 120  
 106 ATCATTTCCGAGATCGGCCAGAACACACGAGCGGACCTGGAGCTAGCCAAAGCGATGATC 165  
 121 CGCATGGCCAAAGAGTGTGGGGCTGATTGTGCCAAGTTCACAGAGAGTGAGTGAATTC 180  
 166 CGCATGGCCAAAGAGTGTGGGGCTGATTGTGCCAAGTTCAGAGAGTGAGTGAATTC 225  
 181 AAGTTTAAATCGAAAGCTTGGAGGCGCATACACTCGAAGCATTCCTGGGGGAAGACG 240  
 226 AAGTTTAAATCGAAAGCTTGGAGGCGCATACACTCGAAGCATTCCTGGGGGAAGACG 285  
 241 TACGGGAGCAACAGCATCTGGAGTTCAGCCATGACGAGTACAGGAGCTGCAGAGG 300  
 286 TACGGGAGCAACAGCATCTGGAGTTCAGCCATGACGAGTACAGGAGCTGCAGAGG 345  
 301 TACGCCGAGGAGTGGAGTCTTCTTCACTGCCTCTGGCATGGATGAGATGGCATTTGAA 360  
 346 TACGCCGAGGAGTGGAGTCTTCTTCACTGCCTCTGGCATGGATGAGATGGCATTTGAA 405  
 361 TTCTGTGATGAATGAATGTTTCAATTTTCAAGTTGGATCTGGAGACACTAATAATTTT 420  
 406 TTCTGTGATGAATGAATGTTTCAATTTTCAAGTTGGATCTGGAGACACTAATAATTTT 465  
 421 CTTTATCTGGAAGAGACAGCCAAAGAGTGCCTCAATGGTGTATCTCAGTGGATGCAG 480  
 466 CTTTATCTGGAAGAGACAGCCAAAGAGTGCCTCAATGGTGTATCTCAGTGGATGCAG 525  
 481 TCAATGGACACCATGAAGCAAGTTTATCAGATCTGTGAAGCCCTCAACCCCACTTCTGC 540  
 526 TCAATGGACACCATGAAGCAAGTTTATCAGATCTGTGAAGCCCTCAACCCCACTTCTGC 585  
 541 TTCTTGAGTGTACAGCGCATACCCGCTCCAGCTCGAGGACGTCACACTGCGGGTCATC 600  
 586 TTCTTGAGTGTACAGCGCATACCCGCTCCAGCTCGAGGACGTCACACTGCGGGTCATC 645

QY 601 TCGGAATATCAGAAGCTCTTTCTCTGACATTCCTCCATGAGGTATTTCTGGGCATGAAACAGGC 660  
 DB 646 TCGGAATATCAGAAGCTCTTTCTCTGACATTCCTCCATGAGGTATTTCTGGGCATGAAACAGGC 705  
 QY 661 ATAGCGATATCTGTCGCGCAGTGGCTCTGGGGCCAAAGTCTTGGAAAGTGCACATAACT 720  
 DB 706 ATAGCGATATCTGTCGCGCAGTGGCTCTGGGGCCAAAGTCTTGGAAAGTGCACATAACT 765  
 QY 721 TTGGACAAGACCTCGAAGGGGAGTGACCACTCGGCTCTGCTGGAGCTTGGAAGTGGCC 780  
 DB 766 TTGGACAAGACCTCGAAGGGGAGTGACCACTCGGCTCTGCTGGAGCTTGGAAGTGGCC 825  
 QY 781 GAGCTGGTGGCTGAGTGGCTCTGAGCGTCTTGTGAGCGTGGCTGGGGTCCCCAAACCAAGCAGCTG 840  
 DB 826 GAGCTGGTGGCTGAGTGGCTCTTGTGAGCGTGGCTGGGGTCCCCAAACCAAGCAGCTG 885  
 QY 841 CTGCCCTGTGAGTGGCTCGAATGAGAAGCTGGGCAAGTCTGTGTGGCCCAAGTGA 900  
 DB 886 CTGCCCTGTGAGTGGCTCGAATGAGAAGCTGGGCAAGTCTGTGTGGCCCAAGTGA 945  
 QY 901 ATTCGGAAGGACCACTTCTTAACAATGACATGCTCACCGTGAAGTGGGTGAGCCCAA 960  
 DB 946 ATTCGGAAGGACCACTTCTTAACAATGACATGCTCACCGTGAAGTGGGTGAGCCCAA 1005  
 QY 961 GCCTATCTCTGGAAGCATCTTTAATCTAGTGGGCAAGAGTCTCTGGTCACTGTTGAA 1020  
 DB 1006 GCCTATCTCTGGAAGCATCTTTAATCTAGTGGGCAAGAGTCTCTGGTCACTGTTGAA 1065  
 QY 1021 GAGATGACACCATCATGGAAGAAATTTGTAGATATATCATGCAAAAATCAAGTCTTAA 1080  
 DB 1066 GAGATGACACCATCATGGAAGAAATTTGTAGATATATCATGCAAAAATCAAGTCTTAA 1125

RESULT 5

ABL90630

ID ABL90630 standard; cDNA; 1268 BP.

XX ABL90630;

XX 24-MAY-2002 (first entry)

XX Human polynucleotide SEQ ID NO 1192.

DE Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
 XX anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anti-ulcer;  
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; human; secreted protein; gene; ss.

OS Homo sapiens.

XX WO200190304-A2.

XX 29-NOV-2001.

XX 18-MAY-2001; 2001WO-US016450.

XX 19-MAY-2000; 2000US-0205515P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2002-122018/16.

XX P-PSDB; ABB90221.

XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and prevention of neural, immune system, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal and proliferative disorders.

PS Claim 4; SEQ ID NO 1192; 2081pp + Sequence Listing; English.

XX The invention relates to novel genes (ABL89449-ABL90853) and proteins  
CC (ABH89040-ABH90444) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing  
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
CC infections. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 1268 BP: 328 A; 304 C; 364 G; 258 T; 0 U; 14 Other

Sequence 1268 BP; 328 A; 304 C; 364 G; 258 T; 0 U; 14 Other;

Query Match 99.8%; Score 1078; DB 6; Length 1268;  
Best Local Similarity 99.8%; Pred. No. 2.1e-314;  
Matches 1078: Conservative 1; Mismatches 1; Indels 0

QY	1	ATGCGCGTGGAGCTGGAGCTGTGTCCCGGCGCTGGGTGGCGGGCGGCAACACCCGTCGCTTC	60
Db	121	ATGCGCGTGGAGCTGGAGCTGTGTCCCGGCGCGCTGGGTGGCGGGCGGCAACACCCGTCGCTTC	180
QY	61	ATCATTGCAGAGATCGGCAGAACACACAGGCGGACCTGGACGTAGTCAAGCGCATGATC	120
Db	181	ATCATTGCAGAGATCGGCAGAACACACAGGCGGACCTGGACGTAGTCAAGCGCATGATC	240
QY	121	CGCATGGCCAAAGAGTGTGGGGCTGATTGTGCCAAAGTTCACAGAGAGTGTAGTGAATTC	180
Db	241	CGCATGGCCAAAGAGTGTGGGGCTGATTGTGCCAAAGTTCACAGAGAGTGTAGTGAATTC	300
QY	181	AAGTTTAAATCGGAAGCCCTCGAGAGGCCATACACCTCGAAGCATTTCTTGGGGGAAGACG	240
Db	301	AAGTTTAAATCGGAAGCCCTTGGAGAGGCCATACACCTCGAAGCATTTCTTGGGGGAAGACG	360
QY	241	TACGGGGAGCAAAACGACATCTGGAGTTCAGCCATGACCAAGTACAGGAGCTGCAGAGG	300
Db	361	TACGGGGAGCAAAACGACATCTGGAGTTCAGCCATGACCAAGTACAGGAGCTGCAGAGG	420
QY	301	TACCGCGAGGAGTTGGGATCTTCTTCACTGCGCTCTGGCATCGATCGATGAGATGCGAGTTGAA	360
Db	421	TACGCGAGGAGGTTGGGATCTTCTTCACTGCGCTCTGGCATCGATCGATGAGATGCGAGTTGAA	480
QY	361	TTCTCTGCATGAACTAATGTTTCCATTTTTCAAAGTTGGATCTTGGAGACACTAATAATTTT	420
Db	481	TTCTCTGCATGAACTAATGTTTCCATTTTTCAAAGTTGGATCTTGGAGACACTAATAATTTT	540
QY	421	CCTTATCTGGAAAAGACAGCCAAAAGGTGCGCCCATGGTGTATCTCCAGTGGGATCGAG	480
Db	541	CCTTATCTGGAAAAGACAGCCAAAAGGTGCGCCCATGGTGTATCTCCAGTGGGATCGAG	600
QY	481	TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTCG	540
Db	601	TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTCG	660
QY	541	TTCTTTGCAGTGTACAGCGGCATACCGGCTCGAGCCTGAGGAGCGTCAA	600
Db	661	TTCTTTGCAGTGTACAGCGGCATACCGGCTCGAGCCTGAGGAGCGTCAA	720
QY	601	TCGGAATATCAGAAGCTCTTTCTGTGACATTCCTCATAGGATTTCTTGGGCATGAACA	660
Db	721	TCGGAATATCAGAAGCTCTTTCTGTGACATTCCTCATAGGATTTCTTGGGCATGAACA	780
QY	661	ATAGCGATATCTGTGGCCGACGTGGCTCTGGGGGGCCAAAGGTGTGGAA	720
Db	781	ATAGCGATATCTGTGGCCGACGTGGCTCTGGGGGGCCAAAGGTGTGGAA	840

Qy	721	TTGGA	CAAGACCTGG	AAAGGGAGTGACCACTCGGCCCTCGCTGAGACCTGGAGAACTGGCC	780
Db	841	TTGGA	CAAGACCTGG	AAAGGGAGTGACCACTCGGCCCTCGCTGAGACCTGGAGAACTGGCC	900
Qy	781	GAGCTGGTGGGGT	CAGTGGGTCTGTGGAGCGTGGCCCTGGGCTCCCCAACCAAGCAGAGCTG	840	
Db	901	GAGCTGGTGGGGT	CAGTGGGTCTGTGGAGCGTGGCCCTGGGCTCCCCAACCAAGCAGAGCTG	960	
Qy	841	CTGCCCTGTGAGATGGCCTGCAATGAGAGCTGGGCAAGTCTCTGGTGGCCAAAGTGAAA	900		
Db	961	CTGCCCTGTGAGATGGCCTGCAATGAGAGCTGGGCAAGTCTCTGGTGGCCAAAGTGAAA	1020		
Qy	901	ATTCGGAAGGCAACCATTTCAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCCAAA	960		
Db	1021	ATTCGGAAGGCAACCATTTCAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCCAAA	1080		
Qy	961	GCCTATCCTCCTGAAAGACATCTTTTAATCTAGTGGGCAAGAAGTCTCTGGTCACTGTTGAA	1020		
Db	1081	GCCTATCCTCCTGAAAGACATCTTTTAATCTAGTGGGCAAGAAGTCTCTGGTCACTGTTGAA	1140		
Qy	1021	GAGGATGACACCATCATGGAAGAAATTGGTAGATAATCATGGCAAAAATAATCAAGTCTTAA	1080		
Db	1141	GAGGATGACACCATCATGGAAGAAATTGGTAGATAATCATGGCAAAAATAATCAAGTCTTAA	1200		

RESULT 6  
AAI59142  
ID AAI59142 standard; cdna; 1230 BP.

## RESULT 6

AAI59142

ID AAI59142 standard; cDNA; 1230 BP.

2  
4  
4  
4  
0  
4  
4  
4  
2 3  
4 3

AC AAI59142:

[illegible]

DT 22-OCT-2001 (first entry)  
XX  
XX Human polynucleotide SEQ ID NO 1345.  
XX  
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

OS Homo sapiens.

2  
1  
5  
4  
2  
5  
2  
  
2  
3  
0  
4  
  
2 X  
0 X

PN WO200153312-A1.

XX  
XX  
TU 2TCCCTO0ZOM NY

26-III.-2001.

FD  
XX  
28-000-2001.

26-DEC-2000: 2000WO-IIS034263-

XX  
FF  
26-DEC-2000; 2000MO-US03534533

XX  
PP 23-DEC-1999. 99JIS-00471275

PR 23=DEC-1999; 990S=004/12/3.  
PB 21=JAN-2000; 2000IS=00488725

PR 21-JAN-2000; 2000US=00488725.  
PB 25-FEB-2000; 2000US=00552317

25-APR-2000; 2000US-00552317.  
PR  
20 JUN 2000  
20 JUN 2000

PR 26-JUN-2000; 2000

PR 19-JUL-2000; 2000US-00620312.  
03 AUG 2000 2000UTC 00052450

PR 03-AUG-2000; 2000US-00653450.  
EZ 14 OCT98 0000  
EZ 14 OCT98 0000

PR 14-SEP-2000; 2000US-00662191.

PR 19-OCT-2000; 2000US-

PR 29-NOV-2000; 2000

XX

PA (HYSE-) HYSEQ INC.

XX The invention relates to human nucleic acids (AAI57798-AAI611369) and the  
CC encoded polypeptides (AAM38642-AAM42213) with neurotropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders. Note: The sequence data for this patent did not form  
CC part of the printed specification  
XX  
SQ Sequence 1230 BP; 310 A; 309 C; 358 G; 253 T; 0 U; 0 Other;

Query Match 99.7%; Score 1076.8; DB 4; Length 1230;  
Best Local Similarity 99.8%; Pred. No. 6.2e-314;  
Matches 1078; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGCGCTGGAGCTGGAGCTGTGTCCTCGGCGCTGGTGGGCGGCAACACCCGTGCTTC 60  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 119 ATGCGCTGGAGCTGGAGCTGTGTCCTCGGCGCTGGTGGGCGGCAACACCCGTGCTTC 178  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 61 ATCATTTGCCGAGATCGGCGGAGACACACAGGCGGACCTGGAGTAGCCAGCGCATGATC 120  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 179 ATCATTTGCCGAGATCGGCGGAGACACACAGGCGGACCTGGAGTAGCCAGCGCATGATC 238  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 121 CGCATGSCCAAGGAGTGGGGCTGATTGTGCAAGTTCCAGAGAGTGGAGTGAATTC 180  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 239 CGCATGSCCAAGGAGTGGGGCTGATTGTGCAAGTTCCAGAGAGTGGAGTGAATTC 298  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 181 AAGTTTAAATCGGAAGCCTTGGAGAGCCATACCTCGAAGCAATTCCTGGGGGAAGACG 240  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 299 AAGTTTAAATCGGAAGCCTTGGAGAGCCATACCTCGAAGCAATTCCTGGGGGAAGACG 358  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 241 TACGGGAGCACAACGACATCTGGAGTTCAGCCATGACAGTACAGGAGCTGCAGAGG 300  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 359 TACGGGAGCACAACGACATCTGGAGTTCAGCCATGACAGTACAGGAGCTGCAGAGG 418  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 301 TACGCCGAGGAGTGGAGTCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 419 TACGCCGAGGAGTGGAGTCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 478  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 361 TTCCTGCATGAATGAATGTTCCATTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 479 TTCCTGCATGAATGAATGTTCCATTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 538  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 421 CCTTATCTGGAAGACAGCCCAAAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 539 CCTTATCTGGAAGACAGCCCAAAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 598  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 481 TCAATGACACCATGAAGCAAGTTTATCAGATCTGGAAGCCCTCAACCCCACTTCTGC 540  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 599 TCAATGACACCATGAAGCAAGTTTATCAGATCTGGAAGCCCTCAACCCCACTTCTGC 658  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 541 TTCTTGAGTGTACAGCGCATACCCGCTCCAGCTTGGAGACGTCACCTGCGGGTCAATC 600  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 659 TTCTTGAGTGTACAGCGCATACCCGCTCCAGCTTGGAGACGTCACCTGCGGGTCAATC 718  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 601 TCGGAATATCAGAGCTTTTCTTGACATTCCTGATAGGTTATCTGGGATGAACAGGC 660  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 719 TCGGAATATCAGAGCTTTTCTTGACATTCCTGATAGGTTATCTGGGATGAACAGGC 778  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 661 ATACCGATATCTGTGGCGGAGTGGCTCTGGGGCCCAAGTGTGGAACTGCATTAAT 720  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 779 ATACCGATATCTGTGGCGGAGTGGCTCTGTGGGGCCCAAGTGTGGAACTGCATTAAT 838  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 721 TTGACAAAGACCTGGAAAGGGAGTGAACCTCTGCGCTGAGGCTGGAGAACTGCCC 780  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 839 TTGGACAAGACCTGGAAAGGGAGTGAACCACTCGGCGCTCGCTGGAGCCTGGAGAACTGGCC 898  
QY 781 GAGCTGGTCCGCTCAGTGGCTCTTGTGGAGCGTGGCTGGCTGCCCAACCAAGCAGCTG 840  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 899 GAGCTGGTCCGCTCAGTGGCTCTTGTGGAGCGTGGCTGCCCAACCAAGCAGCTG 958  
QY 841 CTGCCCCTGTGAGATGGCTGCATGAGAGCTGGGCAAGTCTGTGTGGCCCAAGTGAAA 900  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 959 CTGCCCCTGTGAGATGGCTGCATGAGAGCTGGGCAAGTCTGTGTGGCCCAAGTGAAA 1018  
QY 901 ATTCCGGAAGGCACCATTTCTACATGAGCACTGCTCACCGTGAAGTGGTGGAGCCCA 960  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 1019 ATTCCGGAAGGCACCATTTCTACATGAGCACTGCTCACCGTGAAGTGGTGGAGCCCA 1078  
QY 961 GCCTATCTCTCTGGAAGACATCTTTAATCTAGTGGGCAAGAGTCTCTGTCTACTGTTGA 1020  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 1079 GCCTATCTCTCTGGAAGACATCTTTAATCTAGTGGGCAAGAGTCTCTGTCTACTGTTGA 1138  
QY 1021 GAGGATGACCATCATGGAAGAAATTTGATAGATAATCATGCAAAAAAATCAAGTCTTAA 1080  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 1139 GAGGATGACCATCATGGAAGAAATTTGATAGATAATCATGCAAAAAAATCAAGTCTTAA 1198  
RESULT 7  
ADB49125  
ID ADB49125 standard; cDNA; 1230 BP.  
XX  
AC ADB49125;  
XX AC  
XX 04-DEC-2003 (first entry)  
XX  
DE Novel human cDNA SEQ ID NO 1035.  
XX  
KW ss; cancer; neurodegenerative disease; human.  
XX  
OS Homo sapiens.  
XX  
PN US2003104529-A1.  
XX  
PD 05-JUN-2003.  
XX  
PF 04-JAN-2002; 2002US-00037270.  
XX  
PR 21-JAN-2000; 2000US-00488725.  
PR 25-APR-2000; 2000US-00552317.  
PR 19-JUL-2000; 2000US-00620312.  
XX  
PA (ZHOU/) ZHOU P.  
PA (TANG/) TANG Y T.  
PA (LIUC/) LIU C.  
PA (ASUN/) ASUNDI V.  
PA (DRMA/) DRMANAC R T.  
XX  
PI Zhou P, Tang YT, Liu C, Asundi V, Drmanac RT;  
XX  
DR WPI; 2003-678194/64.  
XX  
PT New polynucleotide, useful for treating diseases e.g., cancer or  
PT neurodegenerative diseases.  
XX  
PS Claim 1; SEQ ID NO 1035; 99pp; English.  
XX  
CC The invention relates to a polynucleotide comprising a sequence given in  
CC the specification, or its mature protein-coding portion, or its  
CC complement. The polynucleotide is useful for treating diseases e.g.,  
CC cancer or neurodegenerative disease and many others listed in the  
CC specification. The present sequence represents a novel human cDNA. Note:  
CC the sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format directly from USPTO  
CC at segdata.uspto.gov/sequence.html?DocID=20030104529.  
XX  
SQ Sequence 1230 BP; 310 A; 309 C; 358 G; 253 T; 0 U; 0 Other;

Query Match		99.7%;	Score 1076.8;	DB 8;	Length 1230;
Best Local Similarity		99.8%;	Pred. No. 6.2e-314;		
Matches 1078;		Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	1	ATGCCGCTGGAGCTGGAGCTGTGTC	CCCGCGCTGGGTGGCGGCAAC	CCCGTGCCTC	60
DB	119	ATGCCGCTGGAGCTGGAGCTGTGTC	CCCGCGCTGGGTGGCGGCAAC	CCCGTGCCTC	178
QY	61	ATCATTCGCCGAGATCGGCAGAAC	CCAGCGGCGACTGGACGTA	CGCAAGCATGATC	120
DB	179	ATCATTCGCCGAGATCGGCAGAAC	CCAGCGGCGACTGGACGTA	CGCAAGCATGATC	238
QY	121	CGCATGGCCAGGAGTGGGGCTGAT	TGTCGAAGTTCGAGAGTGAGT	AGTTC	180
DB	239	CGCATGGCCAGGAGTGGGGCTGAT	TGTCGAAGTTCGAGAGTGAGT	AGTTC	298
QY	181	AAGTTTAATCGAAGCCCTGGAG	AGGCATACACCTCGAAGCTTC	TCTGGGGGAAGACG	240
DB	299	AAGTTTAATCGAAGCCCTGGAG	AGGCATACACCTCGAAGCTTC	TCTGGGGGAAGACG	358
QY	241	TACGGGGAGCAAAACGACATCTG	GAGTTGAGCCATGACCGAGTA	CAGGAGCTGCAGAGG	300
DB	359	TACGGGGAGCAAAACGACATCTG	GAGTTGAGCCATGACCGAGTA	CAGGAGCTGCAGAGG	418
QY	301	TACGCCGAGGAGTTGGGATCTTT	CTCACTGCTCTGGCATGGATG	GATGGCAGTTGAA	360
DB	419	TACGCCGAGGAGTTGGGATCTTT	CTCACTGCTCTGGCATGGATG	GATGGCAGTTGAA	478
QY	361	TTCTTCGATGAATGAATGTTCC	ATTTTCAAAGTTGGATCTGG	GAGACATAATATTTT	420
DB	479	TTCTTCGATGAATGAATGTTCC	ATTTTCAAAGTTGGATCTGG	GAGACATAATATTTT	538
QY	421	CTTATCTGGAAGAGACAGCCAAA	AAAGTCCGCCAATGGTGATCT	CCAGTGGATGCAG	480
DB	539	CTTATCTGGAAGAGACAGCCAAA	AAAGTCCGCCAATGGTGATCT	CCAGTGGATGCAG	598
QY	481	TCAATGGACACCATGAAGCAAG	TTTATCAGATCGTGAAGCCCT	CAACCCCAACTTCTGC	540
DB	599	TCAATGGACACCATGAAGCAAG	TTTATCAGATCGTGAAGCCCT	CAACCCCAACTTCTGC	658
QY	541	TTCTTCGAGTGATACAGCGCAT	ACCCGCTGCCAGCTGAGGAC	GTCAACCTCGGGTTCATC	600
DB	659	TTCTTCGAGTGATACAGCGCAT	ACCCGCTGCCAGCTGAGGAC	GTCAACCTCGGGTTCATC	718
QY	601	TCGGAATATCAGAGCTCTTCTG	ACATCCCATAGGTTATCTGG	GGCATGAACAGGC	660
DB	719	TCGGAATATCAGAGCTCTTCTG	ACATCCCATAGGTTATCTGG	GGCATGAACAGGC	778
QY	661	ATAGCGATATCTGTGCCGAGTG	GGCTCTGGGGCCAGGTTGT	TGGAACTCACATACT	720
DB	779	ATAGCGATATCTGTGCCGAGTG	GGCTCTGGGGCCAGGTTGT	TGGAACTCACATACT	838
QY	721	TTGGCAAGACCTTGAAGGGAGT	GAACCACTCGGCTCTCGAG	CCTGGAGAACTGGCC	780
DB	839	TTGGCAAGACCTTGAAGGGAGT	GAACCACTCGGCTCTCGAG	CCTGGAGAACTGGCC	898
QY	781	GAGCTGGTGGTCACTGTGTG	AGCGTCTCGCTGGCTCCCA	CCAGCAGCTG	840
DB	899	GAGCTGGTGGTCACTGTGTG	AGCGTCTCGCTGGCTCCCA	CCAGCAGCTG	958
QY	841	CTGCCCTGTGATGGCTTCAAT	GAGAGCTGGGCAAGTCTGT	TGGTGGCCAAAGTAAA	900
DB	959	CTGCCCTGTGATGGCTTCAAT	GAGAGCTGGGCAAGTCTGT	TGGTGGCCAAAGTAAA	1018
QY	901	ATTCCGGAAGGCACCATTTTAA	CAATGGACATGCTCACCGT	GAAAGTGGTGAGCCCAA	960
DB	1019	ATTCCGGAAGGCACCATTTTAA	CAATGGACATGCTCACCGT	GAAAGTGGTGAGCCCAA	1078
QY	961	GCCTATCTCTCGAGACATCTT	TAATCTAGTGGGCAAGAGT	CCTCGTCACTGTGAA	1020
DB	1079	GCCTATCTCTCGAGACATCTT	TAATCTAGTGGGCAAGAGT	CCTCGTCACTGTGAA	1138
QY	1021	GAGGATGACACCATCATGGA	AGAAATTGGTAGATAATCAT	TGGCAAAAAAATCAAGTCTTAA	1080

Db	1139	GAGGATGACACCATCATGGA	AGAAATTGGTAGATAATCAT	TGGCAAAAAAATCAAGTCTTAA	1198
RESULT 8					
AAH14557	AAH14557 standard; cDNA; 1170 BP.				
XX	AAH14557;				
AC	26-JUN-2001 (first entry)				
DT	Human cDNA sequence SEQ ID NO:12129.				
DE	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.				
XX	Homo sapiens.				
XX	EP1074617-A2.				
XX	07-FEB-2001.				
XX	28-JUL-2000; 2000EP-00116126.				
PF	29-JUL-1999; 99JP-00248036.				
PR	27-AUG-1999; 99JP-00300253.				
PR	11-JAN-2000; 2000JP-00118776.				
PR	02-MAY-2000; 2000JP-00183767.				
PR	09-JUN-2000; 2000JP-00241899.				
XX	(HELI-) HELIX RES INST.				
PA	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;				
PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;				
PI	WPI; 2001-318749/34.				
DR	Primer sets for synthesising polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.				
XX	Claim 8; SEQ ID NO 12129; 2537pp + Sequence Listing; English.				
XX	The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention				
XX	Sequence 1170 BP; 294 A; 285 C; 347 G; 244 T; 0 U; 0 Other;				
Query Match 99.6%; Score 1075.2; DB 4; Length 1170;					
Best Local Similarity 99.7%; Pred. No. 1.6e-313;					
Matches 1077; Conservative 0; Mismatches 3; Indels 0; Gaps 0;					





Db 69 ATCCGCTGGAGCTGGAGCTGTGTCCGGGCGCTGGGTGGCGGCAACCCGCTGCTC 128  
Qy 61 ATCATTCGCGAGATCGGCCAGAAACAACAGGGCGACCTGGACGTAGCCAGCGCATGATC 120  
Db 129 ATCATTCGCGAGATCGGCCAGAAACAACAGGGCGACCTGGATGTAGCCAGCGCATGATC 188  
Qy 121 CGCATGGCCAGAGGTGGGGCTGATTGTGCCAAGTTCACAGAGAGTGAGCTAGATTTC 180  
Db 189 CGCATGGCCAGAGGTGGGGCTGATTGTGCCAAGTTCACAGAGAGTGAGCTAGATTTC 248  
Qy 181 AAGTTTAAATCGAAAGCCCTCGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240  
Db 249 AAGTTTAAATCGAAAGCCCTCGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 308  
Qy 241 TACGGGGAGCAAAACGACATCTGAGATTAGCCATGACACAGTACAGGAGCTGCAGAGG 300  
Db 309 TACGGGGAGCAAAACGACATCTGAGATTAGCCATGACACAGTACAGGAGCTGCAGAGG 368  
Qy 301 TACGCCGAGGAGGTGGGATCTCTTCACTGCTCTGCGATGGATGAGATGGCAGTTGAA 360  
Db 369 TACGCCGAGGAGGTGGGATCTCTTCACTGCTCTGCGATGGATGAGATGGCAGTTGAA 428  
Qy 361 TTCTGCTGAATGAATGATTCCTTTTCAAGTTGATCTGGAGACACATAAATTTT 420  
Db 429 TTCTGCTGAATGAATGATTCCTTTTCAAGTTGATCTGGAGACACATAAATTTT 488  
Qy 421 CCTATCTGGAAGAGACGCCAAA----- 446  
Db 489 CCTATCTGGAAGAGACGCCAAAAGGGGTGGCACAGTGTTTTAAAGATGTCTGTG 548  
Qy 447 ----- 446  
Db 549 GAGTTCAAGTGAATGATGAGACCAAGTGTGGATGTGGTGGGAGAGTCAGAACCTCTA 608  
Qy 447 -----AGGTGCGCCAAATGGTGAATCTCA 469  
Db 609 AAGAAAGTGTGATGTTGTTGGTGTGATTAATCTCAGGTGCGCCAAATGGTGAATCTCA 668  
Qy 470 GTGGATGAGTCAATGAGACCAATGAGCAAGTTTATCAGATCGTGAAGCCCTCAACC 529  
Db 669 GTGGATGAGTCAATGAGACCAATGAGCAAGTTTATCAGATCGTGAAGCCCTCAACC 728  
Qy 530 CCAACTCTGCTTTGTCAGTGTACAGCGATACCGCTCCAGCTCAGACGTCACACC 589  
Db 729 CCAACTCTGCTTTGTCAGTGTACAGCGATACCGCTCCAGCTCAGACGTCACACC 788  
Qy 590 TGCGGGTCATCTCGGAATATCAGAGCTCTTTTCTGACATTCCTCATAGGGTATCTGGGC 649  
Db 789 TGCGGGTCATCTCGGAATATCAGAGCTCTTTTCTGACATTCCTCATAGGGTATCTGGGC 848  
Qy 650 ATGAACAGGCAATAGGATATCTGTGGCGGAGTGGCTCTGGGGGCGCAAGGTGTTGGAC 709  
Db 849 ATGAACAGGCAATAGGATATCTGTGGCGGAGTGGCTCTGGGGGCGCAAGGTGTTGGAC 908  
Qy 710 GTACATTAATTTTGACAAAGCTCGAAGGGAGTGACCACTCGCGCTCGCTGGAGCCTG 769  
Db 909 GTACATTAATTTTGACAAAGCTCGAAGGGAGTGACCACTCGCGCTCGCTGGAGCCTG 968  
Qy 770 GAGAACTGGCCGAGCTGGTGGCGTCAAGTGCCTTTGTGGAGCGTCCCTGGGCTCCCCAA 829  
Db 969 GAGAACTGGCCGAGCTGGTGGCGTCAAGTGCCTTTGTGGAGCGTCCCTGGGCTCCCCAA 1028  
Qy 830 CCAAGCAGTGTGCTGCTGTGATGCGCTGCAATGAGNAGCTGGCAAGTCTGTGGTG 889  
Db 1029 CCAAGCAGTGTGCTGCTGTGATGCGCTGCAATGAGNAGCTGGCAAGTCTGTGGTG 1088  
Qy 890 CCAAGTGAATTTCCGAGAGGCCATTCATAAATGACATGCTCACCGTGAAGGTGG 949  
Db 1089 CCAAGTGAATTTCCGAGAGGCCATTCATAAATGACATGCTCACCGTGAAGGTGG 1148  
Qy 950 GTGAGCCCAAGCCTATCTCTCTGAAGACATCTTTAATCTAGTGGCAAGAGGCTCTGG 1009  
Db 1149 GTGAGCCCAAGGCTATCTCTCTCTGAAGACATCTTTAATCTAGTGGCAAGAGGCTCTGG 1208

Qy 1010 TCACTGTTGAGAGGATGACACCATCATGGAAGAA 1044  
Db 1209 TCACTGTTGAGAGGATGACACCATCATGGAAGAA 1243  
RESULT 10  
AAS33141/C  
ID AAS33141 standard; cDNA; 1238 BP.  
XX AAS33141;  
AC AAS33141;  
DT 04-DEC-2001 (first entry)  
XX  
DE DNA encoding human secreted protein, Seq ID No 100.  
XX Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;  
KW rheumatoid arthritis; antiarteriosclerotic; cardiac; vascular;  
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;  
KW cytototoxic; Alzheimer's disease; Parkinson's disease; human; cancer;  
KW multiple sclerosis; cancer; hyperproliferative disorder; infection;  
KW Gaucher's disease; neurological disease; cerebrovascular disorder;  
KW thrombosis; wound healing; ss.  
XX Homo sapiens.  
PN WO200155326-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001347.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.

```

PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.

PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-451931/48.
XX P-PSDB; AAU20432.
XX
XX New nucleic acids and polypeptides, useful for diagnosing, preventing or
PT treating medical conditions.
XX
XX Claim 1; SEQ ID NO 100; 753pp; English.
XX
XX The invention relates to novel isolated nucleic acid molecules (I)
CC encoding human secreted proteins (II). (I) and (II) are used to prevent,
CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
CC goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in
CC the prevention, treatment and diagnosis of diseases associated with
CC inappropriate expression of secreted proteins. (I) and complementary
CC sequences may also be used as DNA probes in diagnostic assays (e.g.
CC polymerase chain reactions (PCR)) to detect and quantitate the presence
CC of similar nucleic acid sequences in samples, and so which patients may
CC be in need of restorative therapy. (II) may also be used as antigens in
CC the production of antibodies and in assays to identify modulators
CC (agonists and antagonists) of the expression and activity of the secreted
CC proteins. The anti-(II) antibodies and antagonists may also be used to
CC down regulate expression and activity of (II). The anti-(II) antibodies
CC may also be used as diagnostic agents for detecting the presence of (II)
CC in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The
CC disorders include for example: immune/autoimmune diseases (e.g. HIV
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/
CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and
CC thrombosis), infections caused by bacteria, viruses and fungi and ocular
CC disorders (e.g. corneal infections). (I) and (II), agonists, antagonists
CC and antibodies can also be used to promote wound healing, maintain organs
CC before transplantation, and support cell culture of primary tissues.

Query Match 67.8%; Score 732; DB 4; Length 1238;
Best Local Similarity 99.7%; Pred. No. 6 5e-210;
Matches 732; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 347 AGATGGCAGTTGAATTCCTCATGAATGATGTTCCATTTCCTCAAGTTGGATCTGGAG 406
Db 799 AGATGGCAGTTGAATTCCTCATGAATGATGTTCCATTTCCTCAAGTTGGATCTGGAG 740
QY 407 ACACATAATATTTTCTTATCTCGAAAGACACGCCAAAAGGTGCGCCCAATGGTATCT 466

```

Db 739 ACACATAATATTTTCTTATCTGAAAGAGACAGCCAAAAAGGTCGCCCATGGTGATCT 680  
 Qy 467 CCACTGGGATGAGTCAATGACACCAATGAAGCAAGTTTATCAGATGTAAGCCCTCA 526  
 Db 679 CCACTGGGATGAGTCAATGACACCAATGAAGCAAGTTTATCAGATGTAAGCCCTCA 620  
 Qy 527 ACCCAACTTCTGCTTCTTGGAGTACACAGGATACCCGCTCAGCTCAGGACGCA 586  
 Db 619 ACCCAACTTCTGCTTCTTGGAGTACACAGGATACCCGCTCAGCTCAGGACGCA 560  
 Qy 587 ACCTGGGGTCACTCTGGAAATATCAGAAGCTCTTCTGACATTCCTCATAGGATTTCTG 646  
 Db 559 ACCTGGGGTCACTCTGGAAATATCAGAAGCTCTTCTGACATTCCTCATAGGATTTCTG 500  
 Qy 647 GGCATGAACAGGATAGCATATCTGTGGCGGAGTGGCTCTGGGGGCCAAGGTTGG 706  
 Db 499 GGCATGAACAGGATAGCATATCTGTGGCGGAGTGGCTCTGGGGGCCAAGGTTGG 440  
 Qy 707 AACCTCACAATTTGGACAAGACCTGGAAGGGAGTGACACCTGGCTCGCTGGAGC 766  
 Db 439 AACCTCACAATTTGGACAAGACCTGGAAGGGAGTGACACCTGGCTCGCTGGAGC 380  
 Qy 767 CTGGAGAACTGGCGAGCTGTGGCGTCACTGCGTCTTGTGGAGCGTCCCTGGGCTCC 826  
 Db 379 CTGGAGAACTGGCGAGCTGTGGCGTCACTGCGTCTTGTGGAGCGTCCCTGGGCTCC 320  
 Qy 827 CAACCAAGCAGTGTGCTGCTGTGAGATGGCTTCAATGAGAAGCTGGGCAAGTCTGTGG 886  
 Db 319 CAACCAAGCAGTGTGCTGCTGTGAGATGGCTTCAATGAGAAGCTGGGCAAGTCTGTGG 260  
 Qy 887 TGGCCAAAGTGAATAATTCGGAGGACCACTTCAACATGACATGCTCACCGTGAAGG 946  
 Db 259 TGGCCAAAGTGAATAATTCGGAGGACCACTTCAACATGACATGCTCACCGTGAAGG 200  
 Qy 947 TGGGTGAGCCCAAGGCTATCTCTCTGAAGACATCTTTAATCTAGTGGGCAAGAGGTCC 1006  
 Db 199 TGGGTGAGCCCAAGGCTATCTCTCTGAAGACATCTTTAATCTAGTGGGCAAGAGGTCC 140  
 Qy 1007 TGGTCACCTGTTGAAGAGATGACACCATCATGGAAGATTTGTAGATATCATGCAAAA 1066  
 Db 139 TGGTCACCTGTTGAAGAGATGACACCATCATGGAAGATTTGTAGATATCATGCAAAA 80  
 Qy 1067 AAATCAAGTCTTAA 1080  
 Db 79 AAATCAAGTCTTAA 66  
 RESULT 11  
 ID AAH06847 standard; cDNA; 813 BP.  
 AC AAH06847;  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DE Human cDNA clone (5'-primer) SEQ ID NO:3682.  
 XX  
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000BP-00116126.  
 XX  
 PR 29-JUL-1999; 99JP-00248036.  
 PR 27-AUG-1999; 99JP-00300253.  
 PR 11-JAN-2000; 2000JP-00118776.  
 PR 02-MAY-2000; 2000JP-00183767.  
 PR 09-JUN-2000; 2000JP-00241899.  
 XX

(HELI-) HELIX RES INST.  
 Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 WPI; 2001-318749/34.  
 Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.  
 Claim 1; SEQ ID NO 3682; 2537pp + Sequence Listing; English.  
 The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention  
 SQ Sequence 813 BP; 198 A; 199 C; 243 G; 168 T; 0 U; 5 Other;  
 Query Match 65.8%; Score 710.2; DB 4; Length 813;  
 Best Local Similarity 98.7%; Pred. No. 2e-203;  
 Matches 734; Conservative 0; Mismatches 8; Indels 2; Gaps 2;  
 QY 1 ATGCGCTGGAGCTGGAGCTGTGTCGCGGCGCTGGTGGGGGGGCAACACCCGCTGCTTC 60  
 Db 71 ATGCGCTGGAGCTGGAGCTGTGTCGCGGCGCTGGTGGGGGGGCAACACCCGCTGCTTC 130  
 QY 61 ATCATTTCCGAGATCGGCCAGAACCCAGGGCGACCTGGAGCTGAGCCAGCGCATGATC 120  
 Db 131 ATCATTTCCGAGATCGGCCAGAACCCAGGGCGACCTGGAGCTGAGCCAGCGCATGATC 190  
 QY 121 CGCATGGCCAGGAGTGTGGGGCTGATTTGTGCCAAGTTCCAGAGAGTGAGTGAATTC 180  
 Db 191 CGCATGGCCAGGAGTGTGGGGCTGATTTGTGCCAAGTTCCAGAGAGTGAGTGAATTC 250  
 QY 181 AAGTTTAAATCGAAAAGCCTTGGAGAGGCCATACACTCGAAGCATTCCTGGGGGGAAGCG 240  
 Db 251 AAGTTTAAATCGAAAAGCCTTGGAGAGGCCATACACTCGAAGCATTCCTGGGGGGAAGCG 310  
 QY 241 TACGGGGAGCAAAACGACATCTGAGTTTACGCTATGACCATGACCATGAGGAGCTGCAGAGG 300  
 Db 311 TACGGGGAGCAAAACGACATCTGAGTTTACGCTATGACCATGACCATGAGGAGCTGCAGAGG 370  
 QY 301 TACGGGGAGGAGTTGGGATCTTCTCACTGCTCTGGCATGGATGAGATGCGAGTTGAA 360  
 Db 371 TACGGGGAGGAGTTGGGATCTTCTCACTGCTCTGGCATGGATGAGATGCGAGTTGAA 430  
 QY 361 TTCCTGCGATGAATGAATGTTCCATTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420  
 Db 431 TTCCTGCGATGAATGAATGTTCCATTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 490  
 QY 421 CCTTATCTGGAAGAGACACCCAAAAGGTGCGCCCATGGTGTATCTCCAGTGGATGCGAG 480

Db	491	CCTTATCTGGAAAGACAGCAGNCAAAAAAGGTGCCCAATGGTGATCTCCAGTGGGATGCAG	555
Qy	481	TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTGC	540
Db	551	TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTGC	610
Qy	541	TTCTTGCAGTGTACAGCGCATACCCGCTCCAGCTGAGGACGTCAACTGCGGGTCATC	600
Db	611	TTCTTGCAGTGTACAGCGCATACCCGCTCCAGCTGAGGACGTCAACTGCGGGTCATC	670
Qy	601	TCGGAATATCAGAAGCTCTTTCCCTGCATTTCCCATAGGTTATTTCTGGGCATGAACAGGC	660
Db	671	TCGGAATATCAGAAGCTCTTTCCCTGCATTTCCCATAGGTTATTTCTGGGCATGAACAGGC	730
Qy	661	ATAGCGATATCTGTGGCCGCACTGGC-TCTGGGGCCCAAGGTGTTGGAACGTCACATAAAC	719
Db	731	ATAGCGATATCTGTGGCCGCACTGGC-TCTGGGGCCCAAGGTGTTGGAACGTCACAT-AA	789
Qy	720	TTTGGACAAGCACTTGGAGGGAG	743
Db	790	TTTGGACAAGCACTTGGAGGGAG	813
RESULT 12			
AAH11876/c			
ID	AAH11876 standard; cDNA; 514 BP.		
XX	AAH11876;		
XX	26-JUN-2001 (first entry)		
DT	Human cDNA clone (3'-primer) SEQ ID NO:8711.		
DE	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.		
XX	Homo sapiens.		
KW	EP1074617-A2.		
OS	07-FEB-2001.		
PN	28-JUL-2000; 2000EP-00116126.		
PD	29-JUL-1999; 99JP-00248036.		
XX	27-AUG-1999; 99JP-00300253.		
XX	11-JAN-2000; 2000JP-00118776.		
XX	02-MAY-2000; 2000JP-00183767.		
XX	09-JUN-2000; 2000JP-00241899.		
XX	(HELI-) HELIX RES INST.		
PA	XX		
XX	XX		
PI	Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;		
PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;		
XX	WPI; 2001-318749/34.		
DR	XX		

oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

PD 11-OCT-2001.  
 XX 30-MAR-2001; 2001WO-US008631.  
 XX 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX (HYSE-) HYSEQ INC.  
 PA Drmanac RT, Liu C, Tang YT;  
 XX WPI; 2001-639362/73.  
 XX P-PSDB; ASG27650.  
 DR New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX Claim 1; SEQ ID NO 27641; 103pp; English.  
 PS The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 XX sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping.  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 1417 BP; 353 A; 336 G; 409 G; 319 T; 0 U; 0 Other;  
 Query Match 39.5%; Score 426.4; DB 5; Length 1417;  
 Best Local Similarity 89.4%; Pred. No. 1.2e-117;  
 Matches 588; Conservative 0; Mismatches 46; Indels 24; Gaps 11;  
 447 AGGTGCGCCCAATGGTATCTCCAGTGGGATGAGTCAATGGACACCATGAAGCAAGTTTA 506  
 DB 427 AGGTGCGCCCAATGGTATCTCCAGTGGGATGAGTCAATGGACACCATGAAGCAAGTTAT 486  
 QY 507 TCAGATCGTGAAGCCCTCAACCCCAACTCTGCTTC-TTCAGTGTACAGCGATACC 565  
 DB 487 TCAATCGTGAAGCCCTCAACCCCAACTCTGCTTC-TTCAGTGTACAGCGATACC 546  
 QY 566 CGCTCAGCTGAGACGCTCAACC--TGCGGTCTATCTCGGAATATCAGAACTCTTTCC 623  
 DB 547 CGCTCAGCTGAGACGCTCAACCCCTGCGGGTCTATCTCGGAATATCAGAACTCTTTCC 606  
 QY 624 TGACATTTCC--ATAGGGTATCTGGGCATGAACAGGCATAGCGATATCTGT--GGC 677  
 DB 607 TGACCATTTCCCATATAGGGTATCTGGGCATGAACAGGCATAGCGATATCTTTGGCC 666  
 QY 678 CGCAGTGGCTCTGGGGG-CAAAGGTGTGGAA-----CGTCATAAATCTTGGACAAGAC 731  
 DB 667 CCCAGTGGCTCTGGGGGCCCCAAGGTGTGGGAAGCTCTCAATATAAATTTGGACAAGAC 726  
 QY 732 CTGGAAGGGGAGTG-----ACCACTCGGCTCGCT--GGAGCCTGGAGACTGGCCG-AGC 784  
 DB 727 CGGGGAAGGGGAGTGTAGCCCACTCGCTTGGAGCCTTGGAGAACTGGCCGAGC 786

QY 785 TGGTGGGTGAGTGGCTTCTTGTGGAGCGTCCCTGGGTGCCCAACC-AAGCAGCTGCTG 843  
 DB 787 TGGTGGGTGAGTGGCTTCTTGTGGAGCGTCCCTGGGTGCCCAACCCTAAGCAGCTGCTG 846  
 QY 844 CCCTGTGAGATGGCTGCAA-TGAGAAAGCTGGGCAAGTCTGTGTGGCCAAAGTGAATAAT 902  
 DB 847 CCCTGTGAGATGGCTGCAATTGAGAACTGGGCAAGTCTGTGTGGCCAAAGTGAATAAT 906  
 QY 903 TCCGGAAGGCACCAATTTCAACAATGGACATCTCACCGTGAAGTGGGTGAGCCCAAGC 962  
 DB 907 TCCGGAAGGCACCAATTTCAACAATGGACATCTCACCGTGAAGTGGGTGAGCCCAAGC 966  
 QY 963 CTATCTCTCTGAAGACATCTTTAATCTAGTGGGCAAGAGTCTCTGGTCACTGTGGAAGA 1022  
 DB 967 CTATCTCTCTGAAGACATCTTTAATCTAGTGGGCAAGAGTCTCTGGTCACTGTGGAAGA 1026  
 QY 1023 GGATGACACCATCATGGAAGAATTGGTAGATAATCATGCGCAAAAAATCAAGTCTTTAA 1080  
 DB 1027 GGATGACACCATCATGGAAGAATTGGTAGATAATCATGCGCAAAAAATCAAGTCTTTAA 1084  
 RESULT 14  
 ABN95016/c  
 ID ABN95016 standard; DNA; 464 BP.  
 XX  
 AC ABN95016;  
 XX  
 DT 13-AUG-2002 (first entry)  
 XX  
 DE Gene #1514 used to diagnose liver cancer.  
 XX  
 KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;  
 KW metastatic liver tumour; cytostatic; expression profile; disease state;  
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200229103-A2.  
 XX  
 PD 11-APR-2002.  
 XX  
 PF 02-OCT-2001; 2001WO-US030589.  
 XX  
 PR 02-OCT-2000; 2000US-0237054P.  
 XX  
 PA (GENE-) GENE LOGIC INC.  
 XX  
 PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;  
 XX  
 WPI; 2002-426119/45.  
 DR  
 XX  
 PT Diagnosing and detecting the progression of liver cancer, hepatocellular  
 PT carcinoma or metastatic liver tumor in a patient, involves detecting the  
 PT level of expression of two or more genes in a liver tissue sample.  
 XX  
 PS Claim 1; SEQ ID NO 1514; 298pp; English.  
 XX  
 CC The invention relates to a novel method for diagnosing and detecting the  
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver  
 CC tumor in a patient, and differentiating metastatic liver cancer from  
 CC hepatocellular carcinoma in a patient, involving detecting the level of  
 CC expression of two or more genes represented in ABN93503-ABN97455 in a  
 CC tissue sample. The method of the invention has hepatotropic, and  
 CC cytostatic activity. The method is useful for diagnosing and detecting  
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic  
 CC liver carcinoma in a patient. The method is useful for identifying  
 CC expression profiles which serve as useful diagnostic markers as well as  
 CC markers that can be used to monitor disease states, disease progression,  
 CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX

Search completed: September 16, 2004, 02:09:58  
Job time : 674 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2004, 01:28:31 ; Search time 107 Seconds  
(without alignments)  
5601.373 Million cell updates/sec

Title: US-10-759-277-3

Perfect score: 1080

Sequence: 1 atgcgcgtggagctggagct.....gcaaaaaaatcaagtcttaa 1080

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1080	100.0	1080	4	US-09-516-143A-3
2	1076.8	99.7	1230	4	US-09-620-312D-1035
3	73.6	6.8	1664976	4	US-08-916-421B-1
4	52.6	4.9	1234	4	US-08-976-259-81
5	44.4	4.1	11474	4	US-09-495-406-1
6	43.4	4.0	1083	4	US-09-252-991A-6409
7	43.4	4.0	1104	4	US-09-252-991A-6241
8	41.4	3.8	7218	1	US-08-232-463-14
9	41	3.8	2437	4	US-09-833-381-1478
10	41	3.8	2437	4	US-09-833-381-1479
11	40	3.7	738	4	US-09-252-991A-13238
12	38.2	3.5	363	4	US-09-252-991A-9714
13	38.2	3.5	1254	4	US-09-252-991A-9674
14	38.2	3.5	3654	4	US-09-252-991A-9533
15	36.4	3.4	486	4	US-09-252-991A-13452
16	36.4	3.4	1209	4	US-09-252-991A-13551
17	36.4	3.4	1230	4	US-09-252-991A-13580
18	36.4	3.4	1239	4	US-09-252-991A-13787
19	36	3.3	34094	4	US-09-292-034-1
20	35.6	3.3	642	4	US-09-252-991A-12925
21	35.6	3.3	762	4	US-09-252-991A-13075
22	35.4	3.3	4403765	3	US-09-103-840A-2
23	35.4	3.3	4411529	3	US-09-103-840A-1
24	35	3.2	819	4	US-09-252-991A-15630
25	35	3.2	885	4	US-09-252-991A-15757
26	35	3.2	921	4	US-09-252-991A-15590
27	34.8	3.2	1935	4	US-09-252-991A-8513

```

28 34.8 3.2 1944 4 US-09-252-991A-8364 Sequence 8364, Ap
29 34.8 3.2 2073 4 US-09-252-991A-8332 Sequence 8332, Ap
30 34.6 3.2 627 4 US-09-252-991A-11427 Sequence 11427, A
c 31 34.6 3.2 650 4 US-09-669-751-140 Sequence 140, App
32 34.6 3.2 909 4 US-09-252-991A-11391 Sequence 11391, A
c 33 34.6 3.2 948 4 US-09-252-991A-11445 Sequence 11445, A
34 34.4 3.2 213 4 US-09-930-181-5 Sequence 5, Appli
35 34.4 3.2 2908 4 US-09-930-181-1 Sequence 1, Appli
36 34.4 3.2 3364 4 US-09-930-181-3 Sequence 3, Appli
37 34.2 3.2 690 4 US-09-252-991A-14058 Sequence 14058, A
c 38 34 3.1 4403765 3 US-09-103-840A-2 Sequence 2, Appli
c 39 34 3.1 4411529 3 US-09-103-840A-1 Sequence 1, Appli
c 40 33.8 3.1 399 4 US-09-252-991A-5985 Sequence 5985, Ap
c 41 33.8 3.1 1932 4 US-09-252-991A-5919 Sequence 5919, Ap
c 42 33.8 3.1 2214 4 US-09-252-991A-5968 Sequence 5968, Ap
c 43 33.8 3.1 2370 4 US-09-489-039A-1363 Sequence 1363, Ap
44 33.4 3.1 1756 4 US-09-166-350-30 Sequence 30, Appl
45 33.4 3.1 1756 4 US-09-023-655-954 Sequence 954, App

```

#### ALIGNMENTS

```

RESULT 1
US-09-516-143A-3
; Sequence 3, Application US/09516143A
; Patent No. 6333182
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Glycosylation Enzymes
; FILE REFERENCE: PFS05PCT
; CURRENT APPLICATION NUMBER: US/09/516.143A
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/122,409
; PRIOR FILING DATE: 1999-03-02
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1080
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1077)
US-09-516-143A-3

```

```

Query Match 100.0%; Score 1080; DB 4; Length 1080;
Best Local Similarity 100.0%; Pred. No. 1.3e-302;
Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCCGCTGGAGCTGGAGCTGTGTCGCCGGCGCTGGGTGGCGGGCAACACCCGTCGCTTC 60
Db 1 ATGCCGCTGGAGCTGGAGCTGTGTCGCCGGCGCTGGGTGGCGGGCAACACCCGTCGCTTC 60

Qy 61 ATCATTCGCCAGATCGGCCGCAACACACACCGGCGACCTGACGCTAGCAAGCGCATGATC 120
Db 61 ATCATTCGCCAGATCGGCCGCAACACACACCGGCGACCTGACGCTAGCAAGCGCATGATC 120

Qy 121 CGCATGCCCAAGAGTGTGGGGCTGATTGTGCCAAGTTCAGAAAGAGTGAGCTAGTAATTC 180
Db 121 CGCATGCCCAAGAGTGTGGGGCTGATTGTGCCAAGTTCAGAAAGAGTGAGCTAGTAATTC 180

Qy 181 AAGTTTAAATCGGAAAGCCTTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
Db 181 AAGTTTAAATCGGAAAGCCTTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240

Qy 241 TAGCGGAGACAAAACACATCTGGAGTTCAGCCATGACCAAGTACAGGAGCTGCAGAGG 300
Db 241 TAGCGGAGACAAAACACATCTGGAGTTCAGCCATGACCAAGTACAGGAGCTGCAGAGG 300

Qy 301 TAGCCCGAGAGGTGGGATCTTCTTCACGTGCTCTGGCATGGATGAGATGGCAGTTGAA 360
Db 301 TAGCCCGAGAGGTGGGATCTTCTTCACGTGCTCTGGCATGGATGAGATGGCAGTTGAA 360

```

```

361 TTCTGCGATGAATGAATGTTCCATTTTCAAAGTTGGATCTGGAGACACTAAATATTTT 420
361 TTCTGCGATGAATGAATGTTCCATTTTCAAAGTTGGATCTGGAGACACTAAATATTTT 420
421 CCTTATCTGGAAGACAGACGCAAAAGAGTCCGCCAATGGTGATCTCCAGTGGGATGAG 480
421 CCTTATCTGGAAGACAGACGCAAAAGAGTCCGCCAATGGTGATCTCCAGTGGGATGAG 480
481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTTCAACCCCAACTTCTGC 540
481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTTCAACCCCAACTTCTGC 540
541 TTCTGCGATGACAGGCGATACCCGCTCAGCTGAGGACGTCACCTGCGGTCATC 600
541 TTCTGCGATGACAGGCGATACCCGCTCAGCTGAGGACGTCACCTGCGGTCATC 600
601 TCGGAATATCAGAGCTCTTCTCAGCATTCATAGGCTATCTGGGATGAACAGGC 660
601 TCGGAATATCAGAGCTCTTCTCAGCATTCATAGGCTATCTGGGATGAACAGGC 660
661 ATAGCGATATCTGTGGCGCAGTGGCTCTGGGGCCAAAGGTTTGGAAAGTGCACATAACT 720
661 ATAGCGATATCTGTGGCGCAGTGGCTCTGGGGCCAAAGGTTTGGAAAGTGCACATAACT 720
721 TTGGACAGAGCTGGAAGGGAGTGACCACTCGGCTCGCTGGAGCTGGAGAACTGGCC 780
721 TTGGACAGAGCTGGAAGGGAGTGACCACTCGGCTCGCTGGAGCTGGAGAACTGGCC 780
781 GAGCTGTGGCTCAGTGGCTCTTGTGGAGCTGCTGGGCTCCCAACCAAGCAGCTG 840
781 GAGCTGTGGCTCAGTGGCTCTTGTGGAGCTGCTGGGCTCCCAACCAAGCAGCTG 840
841 CTGCCCTGTGATGCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
841 CTGCCCTGTGATGCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
901 ATTCGGAGGACCATCTTAACATGGACATGCTCAGCGTGAAGTGGGAGCCGAGCCAAA 960
901 ATTCGGAGGACCATCTTAACATGGACATGCTCAGCGTGAAGTGGGAGCCGAGCCAAA 960
961 GCCTATCTCTCTGAGACATCTTTAATCTAGTGGGCAAGAGTCTCTGTCTCACTTTGAA 1020
961 GCCTATCTCTCTGAGACATCTTTAATCTAGTGGGCAAGAGTCTCTGTCTCACTTTGAA 1020
1021 GAGGATGACCATCATGGAAGATTTGGTAGATAATCATGGCAAAATAATCAAGTCTTAA 1080
1021 GAGGATGACCATCATGGAAGATTTGGTAGATAATCATGGCAAAATAATCAAGTCTTAA 1080

```

RESULT 2

```

US-09-620-312D-1035
; Sequence 1035, Application US/09620312D
; Patent No. 656962
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 656962e1 Nucleic Acids and

```

```

; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt FL genes Version 1.0
; SEQ ID NO 1035
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (119)..(1198)
; US-09-620-312D-1035

Query Match 99.7%; Score 1076.8; DB 4; Length 1230;
Best Local Similarity 99.8%; Pred. No. 1.1e-301;
Matches 1078; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGCGGGCAACACCCGTGCTTC 60
Db 119 ATGCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGCGGGCAACACCCGTGCTTC 178
QY 61 ATCATTCGCGAGATCGGCCAGAACACACAGGGCGACCTGGAGCTGAGCAAGCGCATGATC 120
Db 179 ATCATTCGCGAGATCGGCCAGAACACACAGGGCGACCTGGAGCTGAGCAAGCGCATGATC 238
QY 121 CGCATGGCCAAAGAGTGTGGGCTGATTGTCCCAAGTTCAGAGAGTGGAGTGTGAGTTC 180
Db 239 CGCATGGCCAAAGAGTGTGGGCTGATTGTCCCAAGTTCAGAGAGTGGAGTGTGAGTTC 298
QY 181 AAGTTTAACTGGAAGAGCTTGGAGAGGCAATACACCTCGAAGCAATCTCTGGGGGAAGACG 240
Db 299 AAGTTTAACTGGAAGAGCTTGGAGAGGCAATACACCTCGAAGCAATCTCTGGGGGAAGACG 358
QY 241 TACGGGAGCAACAAAGCATCTGGAGTTCAGCCATGACAGTACAGGAGCTGCAGAGG 300
Db 359 TACGGGAGCAACAAAGCATCTGGAGTTCAGCCATGACAGTACAGGAGCTGCAGAGG 418
QY 301 TACGGGAGGAGTGTGGATCTTCTTCACTGCTCTGGCATGGATGAGATGGCACTTGA 360
Db 419 TACGGGAGGAGTGTGGATCTTCTTCACTGCTCTGGCATGGATGAGATGGCACTTGA 478
QY 361 TTCCTGCAATGAATGAATGTTTCCATTTTCAAAGTTGGATCTGGAGACACTAAATATTTT 420
Db 479 TTCCTGCAATGAATGAATGTTTCCATTTTCAAAGTTGGATCTGGAGACACTAAATATTTT 538
QY 421 CCTTATCTGGAAGACAGACGCAAAAGTCCGCCAATGGTGATCTCCAGTGGGATGAG 480
Db 539 CCTTATCTGGAAGACAGACGCAAAAGTCCGCCAATGGTGATCTCCAGTGGGATGAG 598
QY 481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTTCAACCCCAACTTCTGC 540
Db 599 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTTCAACCCCAACTTCTGC 658
QY 541 TTCTTGGAGTGTACAGGCGATACCCGCTCAGCCCTGAGGAGCTCAACTGGGGGTCTATC 600
Db 659 TTCTTGGAGTGTACAGGCGATACCCGCTCAGCCCTGAGGAGCTCAACTGGGGGTCTATC 718
QY 601 TCGGAATATCAGAGCTCTTCTCAGCATTCATAGGCTATCTGGGCAATGAAACAGGC 660
Db 719 TCGGAATATCAGAGCTCTTCTCAGCATTCATAGGCTATCTGGGCAATGAAACAGGC 778
QY 661 ATAGCGATATCTGTGGCGCAGTGGCTCTGGGGCCAAAGGTTTGGAAAGTGCACATAACT 720
Db 779 ATAGCGATATCTGTGGCGCAGTGGCTCTGGGGCCAAAGGTTTGGAAAGTGCACATAACT 838
QY 721 TTGGACAGAGCTGGAAGGGAGTGACCACTCGGCTCGCTGGAGCTGGAGAACTGGCC 780

```

```
Db      839 TTGACAAAGACTGGAAGGGAGTGACCACCTCGCTCGCTGGAGCCCTGGAGAACTGGCC 898
Qy      781 GAGCTGGTGGCGTCACTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCAACCAAGCAGCTG 840
Db      899 GAGCTGGTGGCGTCACTGCGTCTTGTGGAGCGTGCCCTGGGCTCCCAACCAAGCAGCTG 958
Qy      841 CTGCCCTGTGAGATGGCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 900
Db      959 CTGCCCTGTGAGATGGCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAAA 1018
Qy      901 ATTCCGGAAGGACCAATCTTAACAAATGACATGCTCACCGTGAAGGTGGGTGAGCCCAA 960
Db      1019 ATTCCGGAAGGACCAATCTTAACAAATGACATGCTCACCGTGAAGGTGGGTGAGCCCAA 1078
Qy      961 GCCTATCTCTCTGAAGACATCTTTAATCTAGTGGCAGAGAGTCTCTGCTACTGTTGAA 1020
Db      1079 GGTATCTCTCTGAAGACATCTTTAATCTAGTGGCAGAGAGTCTCTGCTACTGTTGAA 1138
Qy      1021 GAGGATGACACCATCATCGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080
Db      1139 GAGGATGACACCATCATCGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1198

RESULT 3
US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6503729
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
```

```
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
```

```
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349491)..(1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1470091)..(1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1569020)..(1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1664854)..(1664855)
; OTHER INFORMATION: n equals a, t, c, or g
```

US-08-916-421B-1

```
Query Match 6.8%; Score 73.6; DB 4; Length 1664976;
Best Local Similarity 46.9%; Pred. No. 9.9e-10;
Matches 300; Conservative 0; Mismatches 334; Indels 6; Gaps 2;

QY 275 ATGACGAGTACAGGAGCTGCAGAGGTACGCCGAGGAGGTGGGATCTTTCACTGCGCT 334
Db 465673 AGGAGGAATCTATGAACTAAAGAAATATGCAGAAAAATTCGAATTATGTTATCTCAA 465732

QY 335 CTGGCATGATGAGATGGAGTTGAATCTCTGCATGAATCTGATTTTCAATTTTCAAG 394
Db 465733 CACCATTAGATTAAATAATGTTGATATATAATAAATGAATGTCCTGCATTAAAA 465792

QY 395 TTGGATCTGGAGACACTAATAATTTTCTTATCTGGAAGACAGCAAAAAAGGTGCGC 454
Db 465793 TTGCTCTGGGTGATTAACTTTTATCCCTTTATTAGAAAAAGTGCAAAACAGGCAAGC 465852

QY 455 CAATGGTGATCTCCAGTGGGATGCAGTCAAT---GGACACCATGAAGCAAGTTTATCAGA 511
```

```
Db 465853 CGGTGATTTTATCTACAGGAATGCTGATATTTGGGAAATTTGGGAAGCAGTTAAGTTT 465912
QY 512 TCGTGAAGCCCTCAACCCCAACTTCTGCTTCTTTCAGTGTACCAAGCATACCCGCTCC 571
Db 465913 TAGAAAAATAATGGATGCAGGATATTTTATTGCAATTGTTATTTTCACTTTACCCAAACC 465972
QY 572 AGCCTGAGGACGTCAACCTGCGGGTCACTCGGAATATCAGAAAGCTTTTCTGCAATTC 631
Db 465973 CTTATGAAGATGTCAATTTAAACGCTATTAAAAACCTTGAAGAATATATTTCAAT---ATCC 466029
QY 632 CCATAGGGTATTCTGGCATGAAAACAGCATAGCATATCTGTGGCCGAGTGGCTCTGG 691
Db 466030 CTGTGGGATATTCTGACCATATCTGGAATACTCGCCCACTAGTTTCTGTGGCTTAG 466089
QY 692 GGGCCAAAGTGTGGAACGTGCATCACTTTTGGCAAGACCTGGGAAGGGAGTGACCACT 751
Db 466090 GAGCGGATGTTATTGAGAAGCACTTTTACCTTAGATAAAAAATATGGAAGTCTCTGATCATG 466149
QY 752 CGGCCTCGCTGGAGCCTGGAGAACTGGCCGAGCTGGTGGTCACTGCGTCTGTGGAGC 811
Db 466150 CTTTGTACAGACCCCAAGAAATTTAAGGAATGTTAATAACATAGATTAGTTGAAA 466209
QY 812 GTGCCCTGGGCTCCCAACCAAGCAGTGTGCTGCCCTGTGAGATGCCCTTGCATGAGAAGC 871
Db 466210 AAATGCTTGGAAAGTGGGAAAAAGATACCAATGCTTCTGAAAGACGCTTATTGTTGAAG 466269
QY 872 TGGGCAAGTCTGTGTGGCCCAAGTGAATAATTCGGAAGG 911
Db 466270 CAAGAAGAAGTATTGTAGCAAAAAGAAATATTAAAAAAGG 466309
```

## RESULT 4

```
US-08-976-259-81
; Sequence 81, Application US/08976259
; Patent No. 6316609
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Choi, Gil H.
; APPLICANT: Weich, Rodney A.
; TITLE OF INVENTION: Nucleotide Sequence of Escherichia coli
; Patent No. 6316609
; NUMBER OF SEQUENCES: 142
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Ave, N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/976,259
; FILING DATE: Herewith
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/031,626 AND US 60/061,953
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0740002/EKS/CBM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 81:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1234 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
```

TOPOLOGY: linear

US-08-976-259-81

Query Match 4.9%; Score 52.6; DB 4; Length 1234;

Best Local Similarity 48.5%; Pred. No. 2.9e-05;

Matches 179; Conservative 0; Mismatches 184; Indels 6; Gaps 1;

QY 62 TCATTGCCGAGATCGCCAGAACACACAGGCGACCTGGACGTAGCCAGCGCATGATCC 121

Db 536 TGATCGCAGAGATTACTACTTCTTTGGTGTATGGACCGATTGAAGTCAATGATTTG 595

QY 122 GCATGCCCAAGAGTGTGGGGCTGATTGTGCCAAAGTTCAGAAAGAGTGGCTAGAAATCA 181

Db 596 TAGCGCCCAAGCAGCAGCGGCTGACTATATCAAACTGCAGAGCGTGTATGTGAAGTT 655

QY 182 AGTTTAAATCGGAAGCCCTTGAGAGGCCATACCTCGAAGCATTCCTGGGGGAAGACGT 241

Db 656 TCTATAGCAGGAGAAAGCTGGAGTCACCGTACCACTC-----TCCTTTTGGCACCACCT 709

QY 242 ACGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCATGACAGGAGCTGCAGAGGT 301

Db 710 TTAGGACTATCGGCATCGCATTTGAATCAATGAAGAGCAATTTTCCTTTGTCGACTCTT 769

QY 302 ACGCCGAGAGGTGGGATCTTCTTCACTGCTCTGGCATGGATGGAGTGGCAGTTGAAT 361

Db 770 TCTGTAAAGAGATTGGTATCGCTGGTTGCTTCTATTTAGATATGCCCTCGTATGAGT 829

QY 362 TCCTGCATGAAGTGTTCATTTTCAAGTTGGATCTGGAGACACTAATAATTTTC 421

Db 830 TCATTGCGCAATTTGAACAGATATGATCAAGCTACCACTATATCTGAACATAAAG 889

QY 422 CTTATCTGG 430

Db 890 ATTATTGG 898

RESULT 5

US-09-495-406-1

; Sequence 1, Application US/09495406

; Patent No. 6503744

; GENERAL INFORMATION:

; APPLICANT: Gilbert, Michel

; APPLICANT: Wakarchuk, Warren W.

; APPLICANT: National Research Council of Canada

; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of

; FILE REFERENCE: 019633-000110US

; CURRENT APPLICATION NUMBER: US/09/495,406

; PRIOR FILING DATE: 2000-01-31

; PRIOR FILING DATE: 1999-02-01

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 1

; LENGTH: 11474

; TYPE: DNA

; ORGANISM: Campylobacter jejuni

; FEATURE:

; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384

; OTHER INFORMATION: including LOS biosynthesis locus

US-09-495-406-1

Query Match

Best Local Similarity 4.1%; Score 44.4; DB 4; Length 11474;

Matches 169; Conservative 0; Mismatches 186; Indels 3; Gaps 1;

QY 363 CCTGCATGAACATGAATGTTCCATTTTTCAGAGTTGGATCTGGAGACACTAATAATTTTC 422

Db 7271 CTTAGAAGATATGGAGATTAGTGCTTTTAAAGTTGGTTTCAGGTGAGTGTAAATAATATCC 7330

QY 423 TTATCTGGAAAAGACAGCCAAAAGGTGCGCCCAATGGTGTATCTCCAGTGGGATGCAGTC 482

Db 7331 GCTTATTAAACACATAGACGCTTTAAAGCCCTATGATAGTAGTACAGCAGGAGTAAATAG 7390

QY 483 AATGACACCATGAAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTGCTT 542

Db 7391 TATTGAAAGTATAAAACCAACTGTAAAAATCTTATTAGACAATGAAATTTCCCTTTGTTT 7450

QY 543 CTTGCAAGTGTACCGCATACCCCTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTT 602

Db 7451 AATGCACACAACCAATCTTTTACCCCAACCCCGCATAACTTTGTAAGATTAAACCGCTATGCT 7510

QY 603 GGAATATCAGAAAGCTCTTTCTCTGACATTCCTAGGGTATTCTGGGCAATGAAACAGGCAAT 662

Db 7511 TGAATTAATAAAGAAATTTTCTGTATGG---TAGGCTTAAGCGACCAACACAGAGATAA 7567

QY 663 AGCGATATCTGTGGCCGAGTGGCTCTGGGGCCAAAGTGTGGAACTGACATAACT 720

Db 7568 TCTTGGCTGTTTAGTGGGTTGCACTTGGTCTTGTGCTTGAAGACATTTTACT 7625

RESULT 6

US-09-252-991A-6409

; Sequence 6409, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-02-18

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 6409

; LENGTH: 1083

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-6409

Query Match

Best Local Similarity 4.0%; Score 43.4; DB 4; Length 1083;

Matches 144; Conservative 0; Mismatches 146; Indels 3; Gaps 1;

QY 26 CCGGGCGCTGGGTGGGGGCAACACCCGCTGCTTTCATCATTCGCGAGATCGCCAGAAC 85

Db 423 CCAAGCGCGTCTGTGGCATCGCCAGGGCGCGAGTATCATCATCAACCGCGTGCAGC 482

QY 86 ACAGGGGCGACCTGGAGTACCGCAAGCGCATGATCCGATGGCCCAAGAGTGTGGGGTG 145

Db 483 CCAATGGCGTGTGTACGGGTGACCCGC---TACCGCTTCGACGAGCAGCGTGGCCTGG 539

QY 146 ATTGTGCCAAGTTCACAGAAGTGTAGATTAAGTTTAACTCGGAAGCCCTGGAGA 205

Db 540 AATCGGCCAGTTTCGCAAGCGCGCGCTTCGAGACCGATCATCTGGCACTGGAAGAG 599

QY 206 GGCCATACCTCGAAGCATTTCTTGGGGGAGAGTACGCGGGAGCAACACGATCTGG 265

Db 600 TCACCACCACTTGTCTCATCCGAGGGAAGCGTTCGAGGTGGTCAAGTACCGACCG 659

QY 266 AGTTACGCCATGACCACTACAGAGCTGAGAGGTACGCGGAGAGGAGGTTGGG 318

Db 660 AACCGTGGGATGCCCACTCAGCCCGCAACTGCTGTAACACTGTGTGTATGGAG 712

RESULT 7

US-09-252-991A-6241

; Sequence 6241, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

```

; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6241
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-6241

Query Match      4.0%;   Score 43.4;   DB 4;   Length 1104;
Best Local Similarity 49.1%;   Pred. No. 0.013;
Matches 144;   Conservative 0;   Mismatches 146;   Indels 3;   Gaps 1;

Qy      26   CCGGGCGCTGGTGGCGGGCAACACCCGCTGCTTCATCATTCGCGAGATCGGCCAGAAC 85
      |||
Db      467   CCAAGCGCGTGTGTGGCATCGCCAGGGCGCGAGTACATCCATCAACGCGGTGCAGC 526

Qy      86   ACAGGGCGGACGTGAGCGTAGCGCAAGCGCATGATCCGCATGGCCAGGAGTGTGGGGCTG 145
      |||
Db      527   CCAATGGCGTGTGTACGGCGTGACCCGC---TACCGCTTCGACGAGCAGCGTGGCCCTGG 583

Qy      146   ATTGTGCCAAGTTCCAGAGAGTGCAGCTAGAAATTCAGTTTAAATCGGAAGCCTTGGAGA 205
      |||
Db      584   AATCGGCCAGTTTTCGCCAAGCGCGCGCTTCGAGACCGATTCATCTGGCACTGGAAGAGG 643

Qy      206   GGCCATACACCTCGAAGCATTCCTGGGGGGAAGAGCTACGGGGGAGCAAAACGCAATCTGG 265
      |||
Db      644   TCACCACCACTTCTGCATCCGAGGGAAAGCGTTCCGAGTGGTCAAGTACCGACCG 703

Qy      266   AGTTCAGCCATGACCAAGTACAGGAGCTGCAGAGGTACGCCGAGAGGTTGGG 318
      |||
Db      704   AACCTGGGATGCCAGCTCAGCCCGCAACTGCTGAACACTGTGTGATGGAG 756

```

RESULT 8  
 US-08-232-463-14/c  
 ; Sequence 14, Application US/08232463  
 ; Patent No. 5670367  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DORNER, F.  
 ; APPLICANT: SCHEIFLINGER, F.  
 ; APPLICANT: FALKNER, F. G.  
 ; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
 ; NUMBER OF SEQUENCES: 52  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley & Lardner  
 ; STREET: 1800 Diagonal Road, Suite 500  
 ; CITY: Alexandria  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22313-0299  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/232,463  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/935,313  
 ; FILING DATE:  
 ; APPLICATION NUMBER: EP 91 114 300.6  
 ; FILING DATE: 26-AUG-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: BENT, Stephen A.  
 ; REGISTRATION NUMBER: 29,768

```

; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-Fls
; US-08-232-463-14

Query Match 3.8%; Score 41.4; DB 1; Length 7218;
Best Local Similarity 4.5%; Pred. No. 0.13;
Matches 18; Conservative 209; Mismatches 170; Indels 0; Gaps 0;

QY 88 CAGGCGGCACTGGACGTGACCAAGCCGATCCGATGCGCAAGAGGTGTGGGGTGAT 147
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1435 CRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1376

QY 148 TGTGCCAAGTTCAGAGAGTGTAGATTCAAGTTTAATCGGAAGCCTTGGAGAGG 207
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1375 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1316

QY 208 CCATACACCTCGAAGCATCTCTGGGGAAGACGTACGGGAGACAAACACATCTGGAG 267
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1315 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1256

QY 268 TTCAGCCATGACCAGTACAGGAGCTGCAGAGTACGCCGAGGAGGTGGGATCTTCTC 327
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1255 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1196

QY 328 ACTGCGCTCTGGCATGGATGAGTGGCAGTTGAATTCCTGCATGAAGTGTCCATTT 387
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1195 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1136

QY 388 TTCAAAGTTGGATCTGGAGACATAATAATTTTCCTTATCTGGAAGACAGCAAAAA 447
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1135 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1076

QY 448 GGTGCGCCCAATGGTGTATCTCCAGTGGGATGCAGTCAA 484
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1075 RRRRRRRRRATCCAGACTCCCTCGACTCAGCCCA 1039

```

```

RESULT 9
US-09-833-381-1478
; Sequence 1478, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1478
; LENGTH: 2437
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2437)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1478

```

Query Match 3.8%; Score 41; DB 4; Length 2437;



	Best Local Similarity	52.7%; Pred. No. 0.095;	Mismatches	89; Conservative	0; Mismatches	80; Indels	0; Gaps	0;
Qy	724	GACAAGACCTTGAAGGGAGTGACCACCTCGCTCGCTGGAGCCTGGAGAATGCGCCGAG	783					
Db	1803	GCCAAGCGGCAGAAAGCACCGCAAGGCCATCGCCCGTTCCAGCAGGCACAGCAGGCGCG	1862					
Qy	784	CTGTGTGGGTCAGTGCCTCTTTGTGGAGGTGTCCTCGGGTCCCCAACCAAGCAGACTGCTGT	843					
Db	1863	CAGGAGCAGCAGGAGGGGCCCATGTGGGCCAGCCCCGAGCCCGACCTGCAGGAGAAG	1922					
Qy	844	CCCTGTGAGATGGCCTGCATGACAAGCTGGGCAAGTCTGTGGTGGCCA	892					
Db	1923	CAGCCTCAGAAGGAGGAGGAGGAGGAGCGCGGAGAGAGTGTGGAGGACA	1971					

```

RESULT 10
US-09-833-381-1479
; Sequence 1479, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robinson, Keith E.
; TITLE OF INVENTION: NO. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1479
; LENGTH: 2437
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2437)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1479

```

	Query Match	3.8%	Score 41;	DB 4;	Length 2437;
	Best Local Similarity	52.7%;	Pred. No. 0.095;		
	Matches	89;	Conservative	0;	Mismatches 80; Indels 0; Gaps 0;
QY	724	GACAAGACCTGGAAAGGGAGTGACCACTCGGCTCGCTGGAGCCTGGAGAACTGGCCGAG	783		
Db	1803	GCCAAAGGGCAGAAGGACCCCAAGGCCATCGCCCGCTTCAGCAGGAGACAGCAGGGCGG	1862		
QY	784	CTGGTGGGTCACTGGCGTCTTGTGGAGCGTGCCTTGGGCTCCCCAACCAAGCAGCTGCTG	843		
Db	1863	CAGGAGCAGCAGGAGGGCCAGTGGGCCCAAGCCCGAGCCGCCACGACTGCAGAGGAAG	1922		
QY	844	CCCTGTGAGATGCCCTGCATATGAGAGCTGGGCAAGTCTGTGTGGCCA	892		
Db	1923	CAGCCTGGAGAAGGAGGAGGAGGAGGGGGGAGAGAGTGTGGAGACA	1971		

RESULT 11  
US-09-252-991A-13238  
; Sequence 13238, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142

```

; SEQ ID NO 13238
; LENGTH: 738
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13238

Query Match          3.7%; Score 40; DB 4; Length 738;
Best Local Similarity 55.9%; Pred. No. 0.1;
Matches 76; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 218 CGAAGCATTCCTCGGGGAAGACGCTACGGGAGACACAAACACATCTGGAGTTCAGCCATG 277
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 317 CGCGCGCGTCTTTGCAGAAACATCTACAAGGAACCTCAACCGGACCTGAACATCCGATTC 376
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 278 ACAAGTACAGGGAGCTGCAGAGTACGCCGAGGAGGTTGGGATCTTCTTCACCTGCCTCTG 337
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 377 CCAACAACGGCTACCTGCAGCGCTGGGCGGAGCAGGCGGTGTGTTGCTCATACCTCCC 436
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 338 GCATGGATGAGATGGC 353
      ||||| ||||| |||||
Db 437 TGACCGTCGAGCAGGC 452
      ||||| ||||| |||||

RESULT 12
US-09-252-991A-9714
; Sequence 9714, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

```

```

RESULT 12
US-09-252-991A-9714
; Sequence 9714, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 9714
; LENGTH: 363
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9714

```

Query Match	3.5%	Score 38.2	DB 4	Length 363
Best Local Similarity	49.3%	Pred. No. 0.23		
Matches 100	Conservative 0	Mismatches 103	Indels 0	Gaps 0
QY	697	AAGGTGTGGAAAGCTGCATATAACTTTTCGACAAGACCTTCGAAGGGAGTGACCACTCGCC	756	
Db	19	ATGATGCCGACGCCTTCGGCGGCTGACCTTGAGCGTTGTGCGAGCAACGCCGGCT	78	
QY	757	TCGCTGAGCCTGGAGAACTTGGCGAGCTGGTTCGGTCAAGTGCCTTTGTGGAGCGTGCC	816	
Db	79	CCGCTGGAGCAGATCGCCCTGACCGAAACAGCGCTGGAGCAGGCCAAGCGGTGGGTGCC	138	
QY	817	CTGGGCTCCCAACCAAGCAGCTGCTCCCTCTGAGATGGCCTGCATAGAAAGCTGGCC	876	
Db	139	ACCGACGACGTGGCGAGCTGAACCTGGCCACGACAAAGTACGCCGGCGCAGATCGC	198	
QY	877	AAGTCTGTGGTGCCCAAAAGTAA	899	
Db	199	ATGACCGCGAGTCTCTACAGAA	221	

RESULT 13  
US-09-252-991A-9674  
; Sequence 9674, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: MICRIFIC ACTS AND

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 9674

; LENGTH: 1254

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-9674

Query Match

Best Local Similarity 3.5%; Score 38.2; DB 4; Length 1254;

Matches 100; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 697 AAGGTGTTGGACGTCACATAACTTTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCC 756

Db 101 ATGATGGCCGAGGCGCTCGCCGCGCTGACCTGACCGGTTGTGCGAGCAACGACCCGGCT 160

QY 757 TCCTGTGAGCCTCGAGAACTGCGCCGAGCTGTGCGTCAAGTGCCTGCTTGTGGAGCGTGCC 816

Db 161 CGCTGGAGCAGATGCGGCTGACCGAACAGAGCGCTGGAGCGGCGTGCGTGC 220

QY 817 CTGGGCTCCCAACCAAGCAGCTGTGCTGCTGTGAGATGCTGCTGCAATGAGAGCTGGGC 876

Db 221 ACCGACGACGTGGCGGAGCTGAACTGGCCCGAGCAAGTACGCGCGCGCAGATCGCG 280

QY 877 AAGTCTGTGTGGCCAAAGTCAA 899

Db 281 ATGACCGGAGTCTCTACAAGAA 303

RESULT 14

US-09-252-991A-9533/c

; Sequence 9533; Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 9533

; LENGTH: 3654

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-9533

Query Match

Best Local Similarity 3.5%; Score 38.2; DB 4; Length 3654;

Matches 100; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 697 AAGGTGTTGGACGTCACATAACTTTGGACAAGACCTGGAAGGGGAGTGACCACTCGGCC 756

Db 550 ATGATGGCCGAGGCGCTCGCCGCGCTGACCTGACGGTTGTGCGAGCAACGACCCGGCT 491

QY 757 TCCTGTGAGCCTCGAGAACTGGCCGAGCTGTGCGTCAAGTGCCTGCTTGTGGAGCGTGCC 816

Db 490 CGCTGGAGCAGATGCGCCTGACCGAACAGAGCGCTGGAGCAGGCCAAAGCGGTGGTGCC 431

QY 817 CTGGGCTCCCAACCAAGCAGCTGTGCTGCTGTGAGATGCTGCTGCAATGAGAGCTGGGC 876

Db 430 ACCGACGACGTGGCGGAGCTGAACTGGCCCGAGCAAGTACGCCGCGCGCAGATCGCG 371

QY 877 AAGTCTGTGTGGCCAAAGTCAA 899

Db 370 ATGACCGCGGAGTCTCTACAAGAA 348

RESULT 15

US-09-252-991A-13452

; Sequence 13452; Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 13452

; LENGTH: 486

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-13452

Query Match

Best Local Similarity 3.4%; Score 36.4; DB 4; Length 486;

Matches 67; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 196 GCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACCTACGGGGAGCACAAA 255

Db 74 GCCTATGTCCGCCAGGCGAGTCGTGCCGGGCTGGCGCTTCCGAGAGCAGGC 133

QY 256 CGACATCTGGAGTTTCAGCCATGACCCAGTACAGGGAGCTGCAGAGGTACGCCGAGGAGG 313

Db 134 CATCAGCAGCAGCGCGGCGAGGCCAGCCAGCGCGATCAGCAGATAGGCGCGCAGGAAG 191

Search completed: September 16, 2004, 05:31:28

Job time : 692 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2004, 04:04:27 : Search time 5258 Seconds  
(without alignments)  
1036.752 Million cell updates/sec

Title: US-10-759-277-3

Perfect score: 1080

Sequence: 1 atgccgctggagctggagct.....gcaaaaaaatcaagtcttaa 1080

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3327077 seqs, 2523723180 residues

Total number of hits satisfying chosen parameters: 6654154

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1080	100.0	1080	9	US-09-984-205-3
2	1080	100.0	1080	9	US-09-930-440B-5
3	1080	100.0	1080	17	US-10-759-277-3
4	1078	99.8	1268	16	US-10-264-237-1192
5	1076.8	99.7	1230	15	US-10-037-270-1035
6	1076.8	99.7	1230	16	US-10-117-722-1035
7	418	38.7	464	9	US-09-880-107-1514
8	371.8	34.4	486	10	US-09-918-995-5220
9	336.8	31.2	442	10	US-09-918-995-28620
10	327.8	30.4	417	10	US-09-918-995-25343
11	315.8	29.2	480	10	US-09-918-995-19353
12	280.6	26.0	404	15	US-10-181-447A-18
13	256.2	23.7	410	15	US-10-106-698-2649
14	180.4	16.7	203	13	US-10-085-783A-31232

15	180.4	16.7	203	16	US-10-242-535A-31232	Sequence 31232, A
16	135	12.5	933	15	US-10-156-761-3360	Sequence 3360, Ap
17	135	12.5	9025608	15	US-10-156-761-1	Sequence 1, Appli
18	101.6	9.4	17276	9	US-09-870-759-83	Sequence 83, Appl
19	101.6	9.4	17276	10	US-09-751-708A-83	Sequence 83, Appl
20	95.2	8.8	25020	16	US-10-132-280-1	Sequence 1, Appli
21	85	7.9	2387	16	US-10-108-260A-1941	Sequence 7, Appli
22	63	5.8	1059	9	US-09-930-440B-7	Sequence 31, Appl
23	62.8	5.8	1000	17	US-10-343-561-31	Sequence 701, App
24	62	5.7	133	9	US-09-925-299-701	Sequence 9075, Ap
25	62	5.7	133	10	US-09-925-299-701	Sequence 701, App
26	60	5.6	60	10	US-09-908-975-9075	Sequence 9075, Ap
27	52.6	4.9	1234	9	US-09-956-004-81	Sequence 81, Appl
28	52.2	4.8	65	10	US-09-908-975-25490	Sequence 25490, A
29	49.6	4.6	408	13	US-10-321-396C-5	Sequence 5, Appli
30	49.6	4.6	579	13	US-10-321-396C-15	Sequence 15, Appl
31	49.6	4.6	618	13	US-10-321-396C-3	Sequence 3, Appli
32	49.6	4.6	650	13	US-10-321-396C-17	Sequence 17, Appl
33	49.6	4.6	677	13	US-10-321-396C-16	Sequence 16, Appl
34	49.6	4.6	692	13	US-10-321-396C-18	Sequence 18, Appl
35	49.6	4.6	828	13	US-10-321-396C-1	Sequence 1, Appli
36	44.4	4.1	11474	9	US-09-816-028A-1	Sequence 1, Appli
37	44.4	4.1	11474	15	US-10-303-161-1	Sequence 1, Appli
38	44.4	4.1	11474	15	US-10-303-118-1	Sequence 1, Appli
39	44.4	4.1	11474	15	US-10-303-128-1	Sequence 1, Appli
40	44.4	4.1	11474	15	US-10-303-134-1	Sequence 1, Appli
41	44.4	4.1	11474	15	US-10-303-162-1	Sequence 1, Appli
42	41	3.8	2437	9	US-09-833-381-1478	Sequence 1478, Ap
43	41	3.8	2437	9	US-09-833-381-1479	Sequence 1479, Ap
44	40.2	3.7	1266	15	US-10-156-761-6436	Sequence 6436, Ap
45	38.8	3.6	473	10	US-09-918-995-26116	Sequence 26116, A

## ALIGNMENTS

## RESULT 1

US-09-984-205-3  
; Sequence 3, Application US/09984205  
; Patent No. US20020137175A1  
; GENERAL INFORMATION:  
; APPLICANT: Coleman, Timothy A. et al.  
; TITLE OF INVENTION: Human Glycosylation Enzymes  
; FILE REFERENCE: PF505D1  
; CURRENT APPLICATION NUMBER: US/09/984,205  
; PRIOR FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: PCT/US00/05325  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 09/516,143  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 60/122,409  
; PRIOR FILING DATE: 1999-03-02  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 1080  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1077)  
US-09-984-205-3

Query Match 100.0%; Score 1080; DB 9; Length 1080;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATCCGCTGGAGTGGAGTGTGTCCCGGCGCTGGTGGCGGCAACACCGTGTTC 60  
Db 1 ATCCGCTGGAGTGGAGTGTGTCCCGGCGCTGGTGGCGGCAACACCGTGTTC 60  
Qy 61 ATCATTCGCCAGATCGCCAGAACCAACGAGGCGACTGGACGTAGCAAGCGCATGATC 120

Db 61 ATCAATGCGAGATCGGCCAGAACACACAGGGGACCTGGACGTAGCCAAAGCGCATGATC 120  
Qy 121 CGCATGGCCAGGAGTGTGGGGCTGATGTGCCAAGTTTCCAGAAAGAGTGAGCTAGAAATTC 180  
Db 121 CGCATGGCCAGGAGTGTGGGGCTGATGTGCCAAGTTTCCAGAAAGAGTGAGCTAGAAATTC 180  
Qy 181 AAGTTTAAATCGAAAGCCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGAGC 240  
Db 181 AAGTTTAAATCGAAAGCCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGAGC 240  
Qy 241 TACGGGAGCACAACAGCATCTGAGTTTCCAGCATGACAGTACAGGAGCTGCAGAGG 300  
Db 241 TACGGGAGCACAACAGCATCTGAGTTTCCAGCATGACAGTACAGGAGCTGCAGAGG 300  
Qy 301 TACGCCGAGGAGTTGGGATCTTTCTACATGCTTGGCATGGATGAGATGGCAGTTGAA 360  
Db 301 TACGCCGAGGAGTTGGGATCTTTCTACATGCTTGGCATGGATGAGATGGCAGTTGAA 360  
Qy 361 TTCTTGCATGAACGAATGTTCCATTTTCAAAGTTTGGATCTGGAGACACTAAATATTTT 420  
Db 361 TTCTTGCATGAACGAATGTTCCATTTTCAAAGTTTGGATCTGGAGACACTAAATATTTT 420  
Qy 421 CCTATCTGGAAGACAGACCAAAAAAGTCCGCCAATGCTGATCTCCAGTGGGATGCAG 480  
Db 421 CCTATCTGGAAGACAGACCAAAAAAGTCCGCCAATGCTGATCTCCAGTGGGATGCAG 480  
Qy 481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTGC 540  
Db 481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTGC 540  
Qy 541 TTCTTGCATGTACAGCGCATACCGCTCCAGCGCTGAGGAGCTCAACCTGGGGTCAATC 600  
Db 541 TTCTTGCATGTACAGCGCATACCGCTCCAGCGCTGAGGAGCTCAACCTGGGGTCAATC 600  
Qy 601 TCGGAATATCAGAAGCTCTTTCTGACATTTCCATAGGTAATCTGGGCATGAACAGGC 660  
Db 601 TCGGAATATCAGAAGCTCTTTCTGACATTTCCATAGGTAATCTGGGCATGAACAGGC 660  
Qy 661 ATAGCGATATCTGTGGCCGAGTGGCTCTGGGGGCCAAGGTTTGGAAAGTCAACATACT 720  
Db 661 ATAGCGATATCTGTGGCCGAGTGGCTCTGGGGGCCAAGGTTTGGAAAGTCAACATACT 720  
Qy 721 TTGGACAGACCTGGAAGGGAGTGACCACTCGGCTCTGAGCGCTGGAGAACTGGCC 780  
Db 721 TTGGACAGACCTGGAAGGGAGTGACCACTCGGCTCTGAGCGCTGGAGAACTGGCC 780  
Qy 781 GAGCTGGTGCAGTGGCTCTTGTGAGCGTGGCCCTGGGCTCCCAACCAAGCAGCTG 840  
Db 781 GAGCTGGTGCAGTGGCTCTTGTGAGCGTGGCCCTGGGCTCCCAACCAAGCAGCTG 840  
Qy 841 CTGGCTGTGAGATGGCCCTGCAATGGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAA 900  
Db 841 CTGGCTGTGAGATGGCCCTGCAATGGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAA 900  
Qy 901 ATTCCGAGGACCACTTCTAAATGAGCATGCTCACCGTGAAGTGGTGGCCCAAA 960  
Db 901 ATTCCGAGGACCACTTCTAAATGAGCATGCTCACCGTGAAGTGGTGGCCCAAA 960  
Qy 961 GCCTATCTCTCTGAAGACATCTTTTATCTAGTGGGCAAGAGTCTCTGGTCACTGTTGAA 1020  
Db 961 GCCTATCTCTCTGAAGACATCTTTTATCTAGTGGGCAAGAGTCTCTGGTCACTGTTGAA 1020  
Qy 1021 GAGGATGACCATCTGAAGATTTGGTAGATATCATGGCAAAAAAATCAAGTCTTAA 1080  
Db 1021 GAGGATGACCATCTGAAGAAATTTGGTAGATATCATGGCAAAAAAATCAAGTCTTAA 1080

RESULT 2

US-09-930-440B-5  
; Sequence 5, Application US/09930440B  
; Patent No. US20020142386A1  
; GENERAL INFORMATION:  
; APPLICANT: Betenbaugh et al.

; TITLE OF INVENTION: Engineering Intracellular Sialylation Pathways  
; FILE REFERENCE: PF509P2  
; CURRENT APPLICATION NUMBER: US/09/930,440B  
; CURRENT FILING DATE: 2001-08-16  
; PRIOR APPLICATION NUMBER: 60/227,579  
; PRIOR FILING DATE: 2000-08-25  
; PRIOR APPLICATION NUMBER: 09/516,793  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 60/169,624  
; PRIOR FILING DATE: 1999-12-08  
; PRIOR APPLICATION NUMBER: 60/122,582  
; PRIOR FILING DATE: 1999-03-02  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 1080  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1080)  
US-09-930-440B-5

Query Match 100.0%; Score 1080; DB 9; Length 1080;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCGCTGGAGCTGGAGCTGTGCCGGGGCGCTGGGTGGGGGGGCAACCCCGTCTTC 60  
Db 1 ATGCGCTGGAGCTGGAGCTGTGCCGGGGCGCTGGGTGGGGGGGCAACCCCGTCTTC 60  
Qy 61 ATCATTCGGAGATCGGCCAGAACACACAGGGCGACCTGGAGCTAGCCAAAGCGCATGATC 120  
Db 61 ATCATTCGGAGATCGGCCAGAACACACAGGGCGACCTGGAGCTAGCCAAAGCGCATGATC 120  
Qy 121 CGCATGGCCAAAGAGTGTGGGGCTGATGTGCCAAGTTTCCAGAAAGAGTGAGCTAGAAATTC 180  
Db 121 CGCATGGCCAAAGAGTGTGGGGCTGATGTGCCAAGTTTCCAGAAAGAGTGAGCTAGAAATTC 180  
Qy 181 AAGTTTAAATCGAAAGCCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGAGC 240  
Db 181 AAGTTTAAATCGAAAGCCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGAGC 240  
Qy 241 TACGGGAGCACAACAGCATCTGAGTTTCCAGCATGACAGTACAGGAGCTGCAGAGG 300  
Db 241 TACGGGAGCACAACAGCATCTGAGTTTCCAGCATGACAGTACAGGAGCTGCAGAGG 300  
Qy 301 TACGCCGAGGAGTTGGGATCTTTCTACATGCTTGGCATGGATGAGATGGCAGTTGAA 360  
Db 301 TACGCCGAGGAGTTGGGATCTTTCTACATGCTTGGCATGGATGAGATGGCAGTTGAA 360  
Qy 361 TTCTTGCATGAACGAATGTTCCATTTTCAAAGTTTGGATCTGGAGACACTAAATATTTT 420  
Db 361 TTCTTGCATGAACGAATGTTCCATTTTCAAAGTTTGGATCTGGAGACACTAAATATTTT 420  
Qy 421 CCTATCTGGAAGACAGACCAAAAAAGTCCGCCAATGCTGATCTCCAGTGGGATGCAG 480  
Db 421 CCTATCTGGAAGACAGACCAAAAAAGTCCGCCAATGCTGATCTCCAGTGGGATGCAG 480  
Qy 481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTGC 540  
Db 481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTGC 540  
Qy 541 TTCTTGCATGTACAGCGCATACCGCTCCAGCGCTGAGGAGCTCAACCTGGGGTCAATC 600  
Db 541 TTCTTGCATGTACAGCGCATACCGCTCCAGCGCTGAGGAGCTCAACCTGGGGTCAATC 600  
Qy 601 TCGGAATATCAGAAGCTCTTTCTGACATTTCCATAGGTAATCTGGGCATGAACAGGC 660  
Db 601 TCGGAATATCAGAAGCTCTTTCTGACATTTCCATAGGTAATCTGGGCATGAACAGGC 660  
Qy 661 ATAGCGATATCTGTGGCCGAGTGGCTCTGGGGGCCAAGGTTTGGAAAGTCAACATACT 720  
Db 661 ATAGCGATATCTGTGGCCGAGTGGCTCTGGGGGCCAAGGTTTGGAAAGTCAACATACT 720

Db 661 ATAGCGATATCTGTGGCCGAGTGGCTCTGGGGGCCAAGGTGTTGGAACGTCACATAACT 720  
QY 721 TTGGACAGACTGGAAGGGAGTACCACCTCGGCTCTGGAGCTTGGAGAACTGGCC 780  
Db 721 TTGGACAGACTGGAAGGGAGTACCACCTCGGCTCTGGAGCTTGGAGAACTGGCC 780  
QY 781 GAGCTGGTGGCTGAGTGGCTCTGTGGAGGCTGCTGGGCTCCGCCAACCAAGCAGCTG 840  
Db 781 GAGCTGGTGGCTGAGTGGCTCTGTGGAGGCTGCTGGGCTCCGCCAACCAAGCAGCTG 840  
QY 841 CTGGCTCTGAGATGGCTGCAATGAGAACTGGGCAAGTCTGTGGTGGCCCAAGTGA 900  
Db 841 CTGGCTCTGAGATGGCTGCAATGAGAACTGGGCAAGTCTGTGGTGGCCCAAGTGA 900  
QY 901 ATTCGGAGGACCACTTCAACATGACATGCTCAACGCTGAAGTGGGTGAGGCCAAA 960  
Db 901 ATTCGGAGGACCACTTCAACATGACATGCTCAACGCTGAAGTGGGTGAGGCCAAA 960  
QY 961 GCCTATCTCTGGAAGACATCTTTAATCTAGTGGGCAAGAGTCTCTGTCACCTGTTGAA 1020  
Db 961 GCCTATCTCTGGAAGACATCTTTAATCTAGTGGGCAAGAGTCTCTGTCACCTGTTGAA 1020  
QY 1021 GAGATGACACCATCATGGAAGAAATGGTAGATAATCATGCAAAAAAATCAAGTCTTAA 1080  
Db 1021 GAGATGACACCATCATGGAAGAAATGGTAGATAATCATGCAAAAAAATCAAGTCTTAA 1080

RESULT 3

US-10-759-277-3  
; Sequence 3, Application US/10759277  
; Publication No. US2004014242A1  
; GENERAL INFORMATION:  
; APPLICANT: Coleman, Timothy A. et al.  
; TITLE OF INVENTION: Human Glycosylation Enzymes  
; FILE REFERENCE: PF505D1  
; CURRENT APPLICATION NUMBER: US/10/759,277  
; CURRENT FILING DATE: 2004-01-20  
; PRIOR FILING DATE: US/09/984,205  
; PRIOR FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: PCT/US00/05325  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 09/516,143  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 60/122,409  
; PRIOR FILING DATE: 1999-03-02  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 1080  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1077)  
US-10-759-277-3

Query Match 100.0%; Score 1080; DB 17; Length 1080;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGCCGCTGGAGCTGGAGCTGTGTCGGGGCGCTGGGTGGGGCGGCAACACCGTGTTC 60  
Db 1 ATGCCGCTGGAGCTGGAGCTGTGTCGGGGCGCTGGGTGGGGCGGCAACACCGTGTTC 60  
QY 61 ATCATTCGGAGATGGCCGAGAACCCACGAGGCGACCTGGACGTAGCCAGCGCATGTC 120  
Db 61 ATCATTCGGAGATGGCCGAGAACCCACGAGGCGACCTGGACGTAGCCAGCGCATGTC 120  
QY 121 CGCATGCCAGGAGTGGGGCTGATGTGCCAAGTTCGAGAGAGTGAATTC 180  
Db 121 CGCATGCCAAGGAGTGGGGCTGATGTGCCAAGTTCGAGAGAGTGAATTC 180  
QY 181 AAGTTTAAATCGAAAGCCTTGGAGAGCCATACACCTCGAAGCATTCTCTGGGGAGACG 240

Db 181 AAGTTTAAATCGAAAGCCTTGGAGAGCCATACACCTCGAAGCATTCCTGGGGAGACG 240  
QY 241 TACGGGAGACAAACGACATCTGGAGTTCAGCCATGACCCAGTACAGGAGCTGCAGAGG 300  
Db 241 TACGGGAGACAAACGACATCTGGAGTTCAGCCATGACCCAGTACAGGAGCTGCAGAGG 300  
QY 301 TACGCCGAGGAGTGGGATCTTCTCACTGCCCTCTGGCATGGATGAGATGGCATTTAA 360  
Db 301 TACGCCGAGGAGTGGGATCTTCTCACTGCCCTCTGGCATGGATGAGATGGCATTTAA 360  
QY 361 TTCCTGCATGAATGAATGTTTCCATTTTCAAGTGGATCTGGAGACACTAATAATTTT 420  
Db 361 TTCCTGCATGAATGAATGTTTCCATTTTCAAGTGGATCTGGAGACACTAATAATTTT 420  
QY 421 CTTTATCTGGAAGAGACAGCCAAAGGTGCCCAATGGTGAATCTCCAGTGGGATGCAG 480  
Db 421 CTTTATCTGGAAGAGACAGCCAAAGGTGCCCAATGGTGAATCTCCAGTGGGATGCAG 480  
QY 481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTTCG 540  
Db 481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTTCG 540  
QY 541 TTCTTGAGTGTACCGGATACCCGCTCCAGCTCGAGGACGTCAACCTGGGGTCACT 600  
Db 541 TTCTTGAGTGTACCGGATACCCGCTCCAGCTCGAGGACGTCAACCTGGGGTCACT 600  
QY 601 TCGGAATATCAGAAAGCTCTTTCTGACATTCCTGAGAGTATTCTGGGCATGAACAGGC 660  
Db 601 TCGGAATATCAGAAAGCTCTTTCTGACATTCCTGAGAGTATTCTGGGCATGAACAGGC 660  
QY 661 ATACGATATCTGTGGCGGAGTGGCTCTGGGGGCCAAGTGTGGAACTGCAATAACT 720  
Db 661 ATACGATATCTGTGGCGGAGTGGCTCTGGGGGCCAAGTGTGGAACTGCAATAACT 720  
QY 721 TTGGACAAGACCTGGAAGGGAGTGACCACTCGGCTCGCTGGAGCTTGGAGAACTGGCC 780  
Db 721 TTGGACAAGACCTGGAAGGGAGTGACCACTCGGCTCGCTGGAGCTTGGAGAACTGGCC 780  
QY 781 GAGCTGTGGCTGAGTGGCTCTTTGTGGAGCTGCTCCCTGGGCTCCCAACCAAGCAGCTG 840  
Db 781 GAGCTGTGGCTGAGTGGCTCTTTGTGGAGCTGCTCCCTGGGCTCCCAACCAAGCAGCTG 840  
QY 841 CTGGCTGTGAGATGGCTGCAATGAGAGCTGGGCAAGTCTGTGGTGGCCAAAGTGA 900  
Db 841 CTGGCTGTGAGATGGCTGCAATGAGAGCTGGGCAAGTCTGTGGTGGCCAAAGTGA 900  
QY 901 ATTCGGAAGGACCACTTCTAAATGACATGCTCAACGCTGAAGTGGGTGAGCCCAA 960  
Db 901 ATTCGGAAGGACCACTTCTAAATGACATGCTCAACGCTGAAGTGGGTGAGCCCAA 960  
QY 961 GCCTATCTCTGGAAGACATCTTTAATCTAGTGGGCAAGAGTCTCTGGTCACTGTTGAA 1020  
Db 961 GCCTATCTCTGGAAGACATCTTTAATCTAGTGGGCAAGAGTCTCTGGTCACTGTTGAA 1020  
QY 1021 GAGATGACACCATCATGGAAGAAATGGTAGATAATCATGCAAAAAAATCAAGTCTTAA 1080  
Db 1021 GAGATGACACCATCATGGAAGAAATGGTAGATAATCATGCAAAAAAATCAAGTCTTAA 1080

RESULT 4

US-10-264-237-1192  
; Sequence 1192, Application US/10264237  
; Publication No. US20040009491A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: FA131P1  
; CURRENT APPLICATION NUMBER: US/10/264,237  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/16450  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: US 60/205,515

```

; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1192
; LENGTH: 1268
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5)..(5)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (17)..(17)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (31)..(31)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (1230)..(1230)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (1257)..(1257)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (1265)..(1265)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-1192

Query Match          99.8%; Score 1078; DB 16; Length 1268;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1078; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1  ATGCGCTGGAGCTGGAGCTGTGTCCTGGGCGCTGGGTGGGGGCAACACCCGNGCTTC 60
Db      121 ATGCGCTGGAGCTGGAGCTGTGTCCTGGGCGCTGGGTGGGGGCAACACCCGNGCTTC 180

Qy      61  ATCATTTGCCGAGATCGGCGCAGAACACACAGGCGGACCTGGAGCTAGCCAAAGCGCATGATC 120
Db      181 ATCATTTGCCGAGATCGGCGCAGAACACACAGGCGGACCTGGAGCTAGCCAAAGCGCATGATC 240

Qy      121 CGCATGGCCAAAGAGTGTGGGCTGATTTGTCGAAGTTCCAGAAAGTGGAGCTAGAAATTC 180
Db      241 CGCATGGCCAAAGAGTGTGGGCTGATTTGTCGAAGTTCCAGAAAGTGGAGCTAGAAATTC 300

Qy      181 AAGTTTAATCGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
Db      301 AAGTTTAATCGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG 360

Qy      241 TACGGGGAGCAACACGACATCTGGAGTTTCAGCCATGACCAAGTACAGGGAGCTGCAGAGG 300
Db      361 TACGGGGAGCAACACGACATCTGGAGTTTCAGCCATGACCAAGTACAGGGAGCTGCAGAGG 420

Qy      301 TACGGCGAGGAGTTGGGATCTTCTCACTGCTTGGCATGGATGAGATGGCAGTTGAA 360
Db      421 TACGGCGAGGAGTTGGGATCTTCTCACTGCTTGGCATGGATGAGATGGCAGTTGAA 480

```

```

US-10-037-270-1035
; Sequence 1035, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Fonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides

```

```

RESULT 5
US-10-037-270-1035
; Sequence 1035, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Fonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides

```

```

; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: PC_FL_genes Version 1.0
; SEQ ID NO 1035
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (119)..(1198)
US-10-037-270-1035

Query Match          99.7%; Score 1076.8; DB 15; Length 1230;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1078; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGCGCTGGAGCTGGAGCTGTGTCGGGGCGCTGGGTGGCGGGCAACACCCGTGCTTC 60
Db 119 ATGCGCTGGAGCTGGAGCTGTGTCGGGGCGCTGGGTGGCGGGCAACACCCGTGCTTC 178

Qy 61 ATCATTCGCGAGATCGGCGAGAACACACAGGGCGACCTGGACGTAGCCAGCGCATGATC 120
Db 179 ATCATTCGCGAGATCGGCGAGAACACACAGGGCGACCTGGACGTAGCCAGCGCATGATC 238

Qy 121 CGCATGCGCAAGAGTGGGGCTGATTGTGCCAAGTTCACAGAGAGTGAAGTTC 180
Db 239 CGCATGCGCAAGAGTGGGGCTGATTGTGCCAAGTTCACAGAGAGTGAAGTTC 298

Qy 181 AAGTTTAAATCGAAAGCTTTGGAGAGCCATACACCTCGAAGCATTTCTGGGGGAAGACG 240
Db 299 AAGTTTAAATCGAAAGCTTTGGAGAGCCATACACCTCGAAGCATTTCTGGGGGAAGACG 358

Qy 241 TACGGGAGCAACACGACATCTGGAGTTGAGCCATGACCGATGACAGGAGCTCGAGG 300
Db 359 TACGGGAGCAACACGACATCTGGAGTTGAGCCATGACCGATGACAGGAGCTCGAGG 418

Qy 301 TACGCCGAGAGGTGGGATCTTCTTCACTGCCCTCTGGCATGGATGAGTGGCGATTGAA 360
Db 419 TACGCCGAGAGGTGGGATCTTCTTCACTGCCCTCTGGCATGGATGAGTGGCGATTGAA 478

Qy 361 TTCCTGATGAATGAATGTTCCATTTTCAAAGTTGGATCTGGAGACACTAAATATTT 420
Db 479 TTCCTGATGAATGAATGTTCCATTTTCAAAGTTGGATCTGGAGACACTAAATATTT 538

Qy 421 CCTTATCTGGAAGAGACAGCCAAAGAGTGGCCCAATGGTATCTCCAGTGGATGAG 480
Db 539 CCTTATCTGGAAGAGACAGCCAAAGAGTGGCCCAATGGTATCTCCAGTGGATGAG 598

Qy 481 TCAATGACACCATGAAGAGTTTATCAGATCTGTGAAGCCCTCAACCCCACTTCTGC 540
Db 599 TCAATGACACCATGAAGAGTTTATCAGATCTGTGAAGCCCTCAACCCCACTTCTGC 658

Qy 541 TTTCTGAGTGTACAGCGCATACCCGCTCCAGCTTCAGGACGTCAACCTCGGGGTATC 600
Db 659 TTTCTGAGTGTACAGCGCATACCCGCTCCAGCTTCAGGACGTCAACCTCGGGGTATC 718

Qy 601 TCGGAATATCAGAGCTCTTCTTGACATTCCTGATAGGTATTTCTGGCATGAACAGGC 660
Db 719 TCGGAATATCAGAGCTCTTCTTGACATTCCTGATAGGTATTTCTGGCATGAACAGGC 778

Qy 661 ATACCGATATCTGTGGCGGAGTGGCTCTGGGGCCCAAGTGTGGAGACGTCACATACT 720
Db 779 ATACCGATATCTGTGGCGGAGTGGCTCTGGGGCCCAAGTGTGGAGACGTCACATACT 838

Qy 721 TTGACAAAGACTTGAAGGGGAGTGACCACTCTCGGCTCGCTGGAGCCCTGGAGAACTGGCC 780
Db 839 TTGACAAAGACTTGAAGGGGAGTGACCACTCTCGGCTCGCTGGAGCCCTGGAGAACTGGCC 898

```

```

Qy 781 GAGCTGTGGCTCAGTGGCTGTGTGGAGCGTCCCTGGCTCCCCACCAAGCAGCTG 840
Db 899 GAGCTGTGGCTCAGTGGCTGTGTGGAGCGTCCCTGGCTCCCCACCAAGCAGCTG 958

Qy 841 CTGCCCTGTGAGATGGCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGA 900
Db 959 CTGCCCTGTGAGATGGCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGA 1018

Qy 901 ATTCGGAAGGCACCATTTCTTAACCAATGGACATGCTCACCGTGAAGTGGGTGAGCCAAA 960
Db 1019 ATTCGGAAGGCACCATTTCTTAACCAATGGACATGCTCACCGTGAAGTGGGTGAGCCAAA 1078

Qy 961 GCCTATCTCTGAAGACATCTTTAATCTAGTGGGCAAGAGTCTGTGCTCAGTGTGAA 1020
Db 1079 GCCTATCTCTGAAGACATCTTTAATCTAGTGGGCAAGAGTCTGTGCTCAGTGTGAA 1138

Qy 1021 GAGCATGACACCATCATGGAAGAAATCGTAGATAATCATGCAAAAAAATCAAGTCTTAA 1080
Db 1139 GAGCATGACACCATCATGGAAGAAATCGTAGATAATCATGCAAAAAAATCAAGTCTTAA 1198

RESULT 6
US-10-117-722-1035
; Sequence 1035, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: PC_FL_genes Version 1.0
; SEQ ID NO 1035
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (119)..(1198)
US-10-117-722-1035

```

```

Query Match          99.7%; Score 1076.8; DB 16; Length 1230;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1078; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGCGCTGGAGCTGGAGCTGTGTCGGGGCGCTGGGTGGCGGGCAACACCCGTGCTTC 60
Db 119 ATGCGCTGGAGCTGGAGCTGTGTCGGGGCGCTGGGTGGCGGGCAACACCCGTGCTTC 178

Qy 61 ATCATTCGCGAGATCGGCGAGAACACACAGGGCGACCTGGACGTAGCCAGCGCATGATC 120
Db 179 ATCATTCGCGAGATCGGCGAGAACACACAGGGCGACCTGGACGTAGCCAGCGCATGATC 238

Qy 121 CGCATGCGCAAGAGTGGGGCTGATTGTGCCAAGTTCACAGAGAGTGAAGTTC 180
Db 239 CGCATGCGCAAGAGTGGGGCTGATTGTGCCAAGTTCACAGAGAGTGAAGTTC 298

Qy 181 AAGTTTAAATCGAAAGCTTTGGAGAGCCATACACCTCGAAGCATTCCTCGGGGAAGACG 240
Db 299 AAGTTTAAATCGAAAGCTTTGGAGAGCCATACACCTCGAAGCATTCCTCGGGGAAGACG 358

```



QY 241 TACGGGAGCACAAACGACATCTGGATTTCAGCCATGACCAAGTACAGGAGCTGCAGAGG 300  
 Db 359 TACGGGAGCACAAACGACATCTGGATTTCAGCCATGACCAAGTACAGGAGCTGCAGAGG 418  
 QY 301 TACGCCGAGGAGTTGGGATCTTCTCACTGCTCTGGCATGGATGAGATGCGAGTTGAA 360  
 Db 419 TACGCCGAGGAGTTGGGATCTTCTCACTGCTCTGGCATGGATGAGATGCGAGTTGAA 478  
 QY 361 TTCTGCGATGAATGAATGTTCCATTTTCAAGTTTGAAGTTGGATCTGGAGACACTAAATTTT 420  
 Db 479 TTCTGCGATGAATGAATGTTCCATTTTCAAGTTTGAAGTTGGATCTGGAGACACTAAATTTT 538  
 QY 421 CCTTATCTGGAAGACAGACCAAAAGGTGCGCAATGCTGATCTCCAGTGGGATGAG 480  
 Db 539 CCTTATCTGGAAGACAGACCAAAAGGTGCGCAATGCTGATCTCCAGTGGGATGAG 598  
 QY 481 TCAATGGACACCATGAAGCAAGTTTATCAGATCTGGAAGCCCTCAACCCCAACTTCTGC 540  
 Db 599 TCAATGGACACCATGAAGCAAGTTTATCAGATCTGGAAGCCCTCAACCCCAACTTCTGC 658  
 QY 541 TTCTTGAGTGTACAGGCGATACCCGCTCCAGCTGAGGAGCTCAACCTGCGGCTCATC 600  
 Db 659 TTCTTGAGTGTACAGGCGATACCCGCTCCAGCTGAGGAGCTCAACCTGCGGCTCATC 718  
 QY 601 TCGGAATATCAGAAGCTCTTTCCTGACATTCCTATAGGCTATCTGGGCAATGAACAGGC 660  
 Db 719 TCGGAATATCAGAAGCTCTTTCCTGACATTCCTATAGGCTATCTGGGCAATGAACAGGC 778  
 QY 661 ATAGCGATATCTGGCCGAGTGGCTCTGGGGCCAAAGTGTGGAACTGCACATAACT 720  
 Db 779 ATAGCGATATCTGGCCGAGTGGCTCTGGGGCCAAAGTGTGGAACTGCACATAACT 838  
 QY 721 TTGACAAAGCTGGAAGGAGTGACCACTCGGCTCGCTGGAGCTTGGAGAACTGGCC 780  
 Db 839 TTGACAAAGCTGGAAGGAGTGACCACTCGGCTCGCTGGAGCTTGGAGAACTGGCC 898  
 QY 781 GAGCTGTGCGGTGAGTGGCTCTGAGAGCTGCTGGGCTCCCAACCAAGCAGCTG 840  
 Db 899 GAGCTGTGCGGTGAGTGGCTCTGAGAGCTGCTGGGCTCCCAACCAAGCAGCTG 958  
 QY 841 CTGCCCTGTGAGTGGCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCCAAGTGAAA 900  
 Db 959 CTGCCCTGTGAGTGGCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCCAAGTGAAA 1018  
 QY 901 ATTCCGAGGACCACTTCAACATGACATGCTCACCCTGAGGCTGGTGGAGCCCAA 960  
 Db 1019 ATTCCGAGGACCACTTCAACATGACATGCTCACCCTGAGGCTGGTGGAGCCCAA 1078  
 QY 961 GCCTATCTCTGAGACATCTTAACTAGTGGGCAAGAGTCTCTGCTCACTGTTGAA 1020  
 Db 1079 GGCTATCTCTGAGACATCTTAACTAGTGGGCAAGAGTCTCTGCTCACTGTTGAA 1138  
 QY 1021 GAGGATGACCACTATGGAAGAATTGGTAGATATCATGCGCAAAAAAATCAAGTCTTAA 1080  
 Db 1139 GAGGATGACCACTATGGAAGAATTGGTAGATATCATGCGCAAAAAAATCAAGTCTTAA 1198

RESULT 7

US-09-880-107-1514/c  
 ; Sequence 1514, Application US/09880107  
 ; Patent No. US20020142981A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Horne, Darci T.  
 ; APPLICANT: Vockley, Joseph G.  
 ; APPLICANT: Scherf, Uwe  
 ; APPLICANT: Gene Logic, Inc.  
 ; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
 ; FILE REFERENCE: 44921-5028-WO  
 ; CURRENT APPLICATION NUMBER: US/09/880,107  
 ; PRIOR FILING DATE: 2001-06-14  
 ; PRIOR APPLICATION NUMBER: US 60/211,379  
 ; PRIOR FILING DATE: 2000-06-14

; PRIOR APPLICATION NUMBER: US 60/237,054  
 ; PRIOR FILING DATE: 2000-10-02  
 ; NUMBER OF SEQ ID NOS: 3950  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1514  
 ; LENGTH: 464  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA621146  
 US-09-880-107-1514

Query March 38.7%; Score 418; DB 9; Length 464;

Best Local Similarity 98.6%; Pred. No. 4.9e-124; Indels 1; Gaps 1;  
 Matches 432; Conservative 0; Mismatches 5;

QY 643 TCTGGCATGAAAACAGGCATAGCATATCTCTGGCCGAGTGGCTCTGGGGCCCAAGGTG 702  
 Db 464 TCTGGCATGAAAACAGGCATAGCATATCTCTGGCCGAGTGGCTCTGGGGCCCAAGGTG 405  
 QY 703 TTGAAAGCTCACATAACTTTTGGACAAGACCTGGAGGGAGTGACCACTCGGCTCGCTG 762  
 Db 404 TTGAAAGCTCACATAACTTTTGGACAAGACCTGGAGGGAGTGACCACTCGGCTCGCTG 345  
 QY 763 GAGCTGTGAGAACTGGCCGAGCTGTCGCGTCACTGTGGAGCGTCTTGTGGAGCGTCCCTGGC 822  
 Db 344 GAGCTGTGAGAACTGGCCGAGCTGTCGCGTCACTGTGGAGCGTCTTGTGGAGCGTCCCTGGC 285  
 QY 823 TCCCAACCAAGCAGCTGCTCCCTGTGAGATGCGCTGCAATGAGAACTGGGCAAGTCT 882  
 Db 284 TCCCAACCAAGCAGCTGCTCCCTGTGAGATGCGCTGCAATGAGAACTGGGCAAGTCT 225  
 QY 883 GTGTGGCCAAAGTGAATAATTCGGAGAGGACCACTTAAACAATGGACATGCTCAACCTG 942  
 Db 224 GTGTGGCCAAAGTGAATAATTCGGAGAGGACCACTTAAACAATGGACATGCTCAACCTG 165  
 QY 943 AAGTGGGTGAGCCCAAGCCTATCTCTCAAGACATCTTTAATCTAGTGGGCAAGAG 1002  
 Db 164 AAGTGGGTGAGCCCAAGCCTATCTCTCAAGACATCTTTAATCTAGTGGGCAAGAG 106  
 QY 1003 GTCTGTGCTACTGTGGAAGAGATGACACCATCATGGAAGAATTGGTAGATAATCATGCG 1062  
 Db 105 GTCTGTGCTACTGTGGAAGAGATGACACCATCATGGAAGAATTGGTAGATAATCATGCG 46  
 QY 1063 AAAAAAATCAAGTCTTAA 1080  
 Db 45 AAAAAAATCAAGTCTTAA 28

RESULT 8

US-09-918-995-5220  
 ; Sequence 5220, Application US/09918995  
 ; Publication No. US20030073623A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hyseq, Inc.  
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
 ; FILE REFERENCE: 20411-756  
 ; CURRENT APPLICATION NUMBER: US/09/918,995  
 ; PRIOR FILING DATE: 2001-07-30  
 ; PRIOR APPLICATION NUMBER: US/09/235,076  
 ; PRIOR FILING DATE: 1999-01-20  
 ; NUMBER OF SEQ ID NOS: 38054  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 5220  
 ; LENGTH: 486  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)...(486)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-918-995-5220

Query Match	34.4%;	Score 371.8;	DB 10;	Length 486;
Best Local Similarity	99.5%;	Pred. NO. 4.3e-109;		
Matches 373;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
Qy	1	ATGCCGCTGGAGCTGGAGCTGTGTCGCGGGCGCTGGGTGGCGGGCAACACCCGTCGCTTC	60	
Db	111	ATGCCGCTGGAGCTGGAGCTGTGTCGCGGGCGCTGGGTGGCGGGCAACACCCGTCGCTTC	170	
Qy	61	ATCATNTGCCGAGATCGGCCAGAAACCAACAGGCGCACTGGACGTAGCCCAAGCGCATGATC	120	
Db	171	ATCATNTGCCGAGATCGGCCAGAAACCAACAGGCGCACTGGATGTAGCCCAAGCGCATGATC	230	
Qy	121	CGCATGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCACAGAGAGTGAGCTAGAAATTC	180	
Db	231	CGCATGCCAAGGAGTGTGGGGCTGATTGTGCCAAGTTCACAGAGAGTGAGCTAGAAATTC	290	
Qy	181	AAGTTTAAATCGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACG	240	
Db	291	AAGTTTAAATCGAAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGNAGACG	350	
Qy	241	TACGGGGAGCAAAACGACATCTGGAGTTCAGCCATGACCAAGTACAGGGAGCTGCAGAGG	300	
Db	351	TACGGGGAGCAAAACGACATCTGGAGTTCAGCCATGACCAAGTACAGGGAGCTGCAGAGG	410	
Qy	301	TACGCCGAGGAGGTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCATGTTCAA	360	
Db	411	TACGCCGAGGAGGTGGGATCTTCTTCACTGCCTCTGGCATGGATGAGATGGCATGTTCAA	470	
Qy	361	TTCCTGCATGAACGTG	375	
Db	471	TTCCTGCATGAACGTG	485	

```

RESULT 9
US-09-918-995-28620
; Sequence 28620, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28620
; LENGTH: 442
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(442)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-28620

Query Match          31.2%; Score 336.8; DB 10; Length 442;
Best Local Similarity 98.0%; Pred. No. 8.5e-98;
Matches 341; Conservative 0; Mismatches 7; Indels 0; Gaps 0

QY      2   TGC CGCTGG AGCTGG AGCTGTGTCCCGGCGCTGGTGGGCGGCAACACCGTGTCTTCA 61
Db      90   TGC CGTTTTT AGCTGG AGCTGAGTCCC CGGCGCTGGTGGGCGGCAACACCGTGTCTTGA 149

QY      62   TCATTGCC GAGATCGG CCGAGAAC CACAGAGGCGACCTGGAGCTAGCCAAAGCGCATGTACC 121
Db      150  TCATTGCC GAGATCGG CCGAGAAC CACAGAGGCGACCTGGAGCTAGCCAAAGCGCATGTACC 209

QY      122  GCATGGCC AAGGAGTGTGGGCTGATTGTG CCAAGTTTCCAGGAAGAGTGAGCTAGAAATTC A 181

```

```

Db      210  GCATGGCCAGAGTGTGGGCTGATTGTGCTAAGTTCGAGAAGAGTGCAGCTAGAATTCA  269
QY      182  AGTTTAATCGAAAGCCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACGT  241
Db      270  AGTTTAATCGAAAGCCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACGT  329
QY      242  ACGGGGAGCACAAACGACATCTGAGTTCAGCCATGACCAGTACAGGAGCTGCAGAGGT  301
Db      330  ACGGGGAGCACAAACGACATCTGGAGTTCAGCCATGACCAGTACAGGAGCTGCAGAGGT  389
QY      302  ACGCCGAGGAGGTTGGGATCTTCTTCACTGCCTCTGGCATCGATGAGA  349
Db      390  ACGCCGAGGAGGCTGGGATCTTCTTCACTGCCTCTGGCATCGATGAGA  437

RESULT 10
US-09-918-995-25343
; Sequence 25343, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25343
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(417)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-25343

```

RESULT 11  
US-09-918-995-19353  
; Sequence 19353, Application US/09918995  
; Publication No. US20030073623A1

```

; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19353
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(480)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-19353

Query Match      29.2%; Score 315.8; DB 10; Length 480;
Best Local Similarity 99.4%; Pred. No. 5.4e-91;
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 750 CTCGGCTCGTGGAGCTGAGAACTGGCCGAGCTGGTGGCTCAGTGGCTTTGTGGA 809
DB 51 CTCGGCTCGTGGAGCTGAGAACTGGCCGAGCTGGTGGCTCAGTGGCTTTGTGGA 110

QY 810 GCGTGCCCTGGCTCCCAACCAACAGCAGCTGCTGCCCTGTGAGATGGCTGCAATGAGAA 869
DB 111 GCGTGCCCTGGCTCCCAACCAACAGCAGCTGCTGCCCTGTGAGATGGCTGCAATGAGAA 170

QY 870 GCTGGCAAGTCTGTGGTGGCCAAAGTGAAATTCGGGAAGCACAATCTTAACAATGGA 929
DB 171 GCTGGCAAGTCTGTGGTGGCCAAAGTGAAATTCGGGAAGCACAATCTTAACAATGGA 230

QY 930 CATGCTACCGTGAAGTGGTGAGCCCAAGCCCTATCTCTGAAGACATCTTAATCT 989
DB 231 CATGCTACCGTGAAGTGGTGAGCCCAAGCCCTATCTCTGAAGACATCTTAATCT 290

QY 990 AGTGGGCAAGAGTCTGCTGGTCACTGTTGAAGAGGATGACACCATCATGGAAGAAATGGT 1049
DB 291 AGTGGGCAAGAGTCTGCTGGTCACTGTTGAAGAGGATGACACCATCATGGAAGAAATGGT 350

QY 1050 AGATAATCATGCAAAAAA 1068
DB 351 AGATAATCATGCAAAAAA 369

```

```

RESULT 12
US-10-181-447A-18
; Sequence 18, Application US/10181447A
; Publication No. US20030180738A1
; GENERAL INFORMATION:
; APPLICANT: The No. US20030180738A1Althingham Trent University
; TITLE OF INVENTION: Cancer Associated Genes and their Products
; FILE REFERENCE: No. US20030180738A1Althingham Trent Uni
; CURRENT APPLICATION NUMBER: US/10/181,447A
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/GB/01/000188
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: GB0000993.6
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 404
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(404)

```

```

; OTHER INFORMATION: n = g, a, t, or c
US-10-181-447A-18

Query Match      26.0%; Score 280.6; DB 15; Length 404;
Best Local Similarity 98.3%; Pred. No. 1.2e-79;
Matches 283; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGCGCTGGAGCTGGAGCTGTCTCCGGGCGCTGGGTGGGGGGGCAACCCGTCGCTTC 60
DB 104 ATGCGCTGGAGCTGGAGCTGTCTCCGGGCGCTGGGTGGGGGGGCAACCCGTCGCTTC 163

QY 61 ATCATTGCGGAGATCGGCCAGAACACACAGGCGGACCTGGACGCTAGCCAAAGCGCATGATC 120
DB 164 ATCATTGCGGAGATCGGCCAGAACACACAGGCGGACCTGGACGCTAGCCAAAGCGCATGATC 223

QY 121 CGCATGGCCAAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAGAGTGAGCTAGAAATTC 180
DB 224 CGCATGGCCAAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAGAGTGAGCTAGAAATTC 283

QY 181 AAGTTTAAATCGAAAGCCTTGGAGAGGCGCATACACCTCGAAGCATTCCTGGGGGAAGACG 240
DB 284 AAGTTTAAATCGAAAGCCTTGGAGAGGCGCATACACCTCGAAGCATTCCTGGGGGAAGACG 343

QY 241 TACGGGGAGCACAAACGACATCTGGAGTTTCAGCCATGACCATACAGG 288
DB 344 TACGGGGAGCACAAACGACATCTGGAGTTTCAGCCATGACCATACAGG 391

```

```

RESULT 13
US-10-106-698-2649
; Sequence 2649, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005PI
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 2649
; LENGTH: 410
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (287)..(287)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (300)..(300)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (370)..(370)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (374)..(374)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (402)..(402)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (405)..(405)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (408)..(408)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-2649

```

Query Match 23.7%; Score 256.2; DB 15; Length 410;  
Best Local Similarity 97.2%; Pred. No. 9.2e-72;  
Matches 280; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

Qy 548 AGTGATACAGCGCATACCGCTCCAGCTGAGGAGCTCAACCTGGGGTCTATCTCGGAAT 607  
|||||  
Db 110 AGTGATACAGCGCATACCGCTCCAGCTGAGGAGCTCAACCTGGGGTCTATCTCGGAAT 169  
|||||

Qy 608 ATCAGAAGCTCTTCTGACATTCCTATAGGATATCTGGGATGAAACAGGCATAGCGA 667  
|||||  
Db 170 ATCAGAAGCTCTTCTGACATTCCTATAGGATATCTGGGATGAAACAGGCATAGCGA 229  
|||||

Qy 668 TATCTGTGGCGCAGTGGCTCTGGGGCCAGAGTGTGGAAGCTCACATACTTTGGACA 727  
|||||  
Db 230 TATCTGTGGCGCAGTGGCTCTGGGGCCAGAGTGTGGAAGCTCACATACTTTGGNCA 289  
|||||

Qy 728 AGACCTGGAAGGGAGTGACCACTCGGCTCGCTGGAGCCTGGAGAACTGGCCG-AGCTG 786  
|||||  
Db 290 AGACCTGGAAGGGAGTGACCACTCGGACTCGCTGGAGCCTGGAGAACTGGGCGAAGCTG 349  
|||||

Qy 787 GTGGGTCAAGTGGTCTTGT-GGAGCGTGCCTTGGGCTCCGCCAACCAA 833  
|||||  
Db 350 GTGGGTCAAGTGGTCTTGTGGAGCGTGCCTTGGGCTCCGCCAACCAA 397  
|||||

RESULT 14  
US-10-085-783A-31232  
; Sequence 31232, Application US/10085783A  
; Publication No. US20040037841A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2002  
; CURRENT APPLICATION NUMBER: US/10/085,783A  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 31232  
; LENGTH: 203  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (6)..(6)  
; OTHER INFORMATION: n is a, c, g, or t  
; NAME/KEY: misc feature  
; LOCATION: (31)..(31)  
; OTHER INFORMATION: n is a, c, g, or t  
US-10-085-783A-31232

Query Match 16.7%; Score 180.4; DB 13; Length 203;  
Best Local Similarity 98.4%; Pred. No. 1.9e-47;  
Matches 181; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 897 GAAATTCGGGAGGACCATCTTAAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCC 956  
|||||  
Db 1 GAAATTCGGGAGGACCATCTTAAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCC 60  
|||||

Qy 957 CAAAGCCTATCCTCTGGAAGACATCTTTAATCTAGTGGGCAAGAGGTCTCTGGTCACTGT 1016  
|||||  
Db 61 CAAAGGCTATCCTCTGGAAGACATCTTTAATCTAGTGGGCAAGAGGTCTCTGGTCACTGT 120  
|||||

Qy 1017 TGAAGAGGATGACACCATCATGGAAGAAATTGGTAGAATATCATGCGCAAAAAATCAAGTC 1076  
|||||

Db 121 TGAAGAGGATGACACCATCATGGAAGAAATTGGTAGAATATCATGCGCAAAAAATCAAGTC 180  
Qy 1077 TTAA 1080  
|||||  
Db 181 TTAA 184  
|||||

RESULT 15  
US-10-242-535A-31232  
; Sequence 31232, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2005  
; CURRENT APPLICATION NUMBER: US/10/242,535A  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 10/085,783  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 31232  
; LENGTH: 203  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (6)..(6)  
; OTHER INFORMATION: n is a, c, g, or t  
; NAME/KEY: misc feature  
; LOCATION: (31)..(31)  
; OTHER INFORMATION: n is a, c, g, or t  
US-10-242-535A-31232

Query Match 16.7%; Score 180.4; DB 16; Length 203;  
Best Local Similarity 98.4%; Pred. No. 1.9e-47;  
Matches 181; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 897 GAAATTCGGGAGGACCATCTTAAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCC 956  
|||||  
Db 1 GAAATTCGGGAGGACCATCTTAAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCC 60  
|||||

Qy 957 CAAAGCCTATCCTCTGGAAGACATCTTTAATCTAGTGGGCAAGAGGTCTCTGGTCACTGT 1016  
|||||  
Db 61 CAAAGGCTATCCTCTGGAAGACATCTTTAATCTAGTGGGCAAGAGGTCTCTGGTCACTGT 120  
|||||

Qy 1017 TGAAGAGGATGACACCATCATGGAAGAAATTGGTAGAATATCATGCGCAAAAAATCAAGTC 1076  
|||||  
Db 121 TGAAGAGGATGACACCATCATGGAAGAAATTGGTAGAATATCATGCGCAAAAAATCAAGTC 180  
Qy 1077 TTAA 1080  
|||||  
Db 181 TTAA 184  
|||||

Search completed: September 16, 2004, 09:36:47  
Job time : 5264 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2004, 01:01:16 ; Search time 4417 Seconds  
(without alignments)  
7301.597 Million cell updates/sec

Title: US-10-759-277-3

Perfect score: 1080

Sequence: 1 atgcgcgtgagctggagct.....gcaaaaaaatcaagctcttaa 1080

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1076.8	99.7	1080	29	AY414837	AY414837 Homo sapi
C 2	952.6	88.2	984	9	AL578658	AL578658 AL578658
C 3	949.6	87.9	1201	13	BX384733	BX384733 BX384733
C 4	917.8	85.0	990	9	AL578579	AL578579 AL578579

5	883.2	81.8	1080	29	AY414839	AY414839 Mus muscu
6	883.2	81.8	1906	11	AK076290	AK076290 Mus muscu
7	870.6	80.6	2109	11	AK088912	AK088912 Mus muscu
8	856.6	79.3	925	12	BI517350	BI517350 603041586
9	856.4	79.3	987	9	AL556387	AL556387 AL556387
10	843	78.1	1005	9	AL556278	AL556278 AL556278
11	840	77.8	1201	9	AL515470	AL515470 AL515470
12	830	76.9	989	9	AL526241	AL526241 AL526241
13	823.4	76.2	889	13	BU187658	BU187658 AGENCOURT
14	803.2	74.4	872	13	BQ432284	BQ432284 AGENCOURT
15	796.4	73.7	1031	13	BQ064521	BQ064521 AGENCOURT
16	795.8	73.7	842	9	AL522851	AL522851 AL522851
17	792.2	73.4	1049	12	BM541313	BM541313 AGENCOURT
18	789.4	73.1	948	29	AY414838	AY414838 Pan trogl
19	786	72.8	1046	9	AL526281	AL526281 AL526281
20	784.4	72.6	890	13	BU181010	BU181010 AGENCOURT
21	775	71.8	979	12	BG681730	BG681730 602627851
22	772	71.5	905	13	BQ893527	BQ893527 AGENCOURT
23	764.4	70.8	1097	12	BM806099	BM806099 AGENCOURT
24	757.4	70.1	878	9	AL522852	AL522852 AL522852
25	754.4	69.9	878	12	BI824205	BI824205 603040573
26	749.2	69.4	1070	9	AL544111	AL544111 AL544111
27	737.6	68.3	1050	12	BM563430	BM563430 AGENCOURT
28	736.2	68.2	773	12	BI517963	BI517963 603041586
29	735.4	68.1	845	13	BU528624	BU528624 AGENCOURT
30	726.2	67.2	905	13	BQ649507	BQ649507 AGENCOURT
31	722.2	66.9	883	12	BI258475	BI258475 602972467
32	721.6	66.8	777	12	BG568099	BG568099 602586980
33	717.4	66.4	978	12	BG114903	BG114903 602315105
34	710.6	65.8	919	13	BU195376	BU195376 AGENCOURT
35	710.2	65.8	813	9	AU132828	AU132828 AU132828
36	703.2	65.1	1031	9	AL544197	AL544197 AL544197
37	695	64.4	706	12	BG746292	BG746292 602703522
38	690	63.9	690	12	BG746329	BG746329 602703579
39	689.4	63.8	835	10	BE614630	BE614630 601504568
40	687.8	63.7	948	10	BE793764	BE793764 601589915
41	686.4	63.6	769	13	BU618375	BU618375 UI-H-PH1 -
42	682.2	63.2	779	14	CA417480	CA417480 UI-H-PE0 -
43	680.4	63.0	931	12	BI525002	BI525002 602924077
44	678.4	62.8	844	13	BU167760	BU167760 AGENCOURT
45	675.2	62.5	877	12	BI905036	BI905036 603169320

## ALIGNMENTS

RESULT 1	AY414837	1080 bp	DNA	linear	GSS 17-DEC-2003
LOCUS	AY414837	Homo sapiens NANS gene, VIRTUAL TRANSCRIPT, partial sequence,			
DEFINITION	AY414837	Genomic survey sequence.			
ACCESSION	AY414837.1	GI:39770796			
VERSION	AY414837.1	GI:39770796			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Infering nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCE	2 (bases 1 to 1080)				
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Direct Submission				

Qy	841	CTGCCCTGTGAGATGGCCTGCAATGAGAACTGGGCAAGTCTGTGTGGCCAAAGTCAA	900
D6	841	CTGCCCTGTGAGATGGCCTGCAATGAGAACTGGGCAAGTCTGTGTGGCCAAAGTCAA	900

Db	841	CTGCCCCGTGAGATGGCCCTTCGAATGAGAAGCTGGCAAGTCTCTGGTGGGCCAAGTGA	900
Qy	901	ATTCCGGAAGGCCACCATTTCTAAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA	960
Db	901	ATTCCGGAAGGCCACCATTTCTAAACAATGGACATGCTCACCGTGAAGGTGGGTGAGCCCAAA	960
Qy	961	GCCTATCCTCCTGAAGACATCTTTTAATCTAGTGGGGAAGAAGTCTCTGGTCACTGTGAA	1020
Db	961	GGCTATCCTCCTGAAGACATCTTTTAATCTAGTGGGGAAGAAGTCTCTGGTCACTGTGAA	1020

Accession	Sequence	Length
D5	1 021 GAGGATGACACCATCATGGGAATTCGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA	1 080
QY	1 021 GAGGATGACACCATCATGGGAATTCGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA	1 080
D10	1 021 GAGGATGACACCATCATGGGAATTCGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA	1 080

RESULT 2	AL578658	Homo sapiens HELA CELLS	984 bp	mRNA	linear	EST 01-JUN-2003
LOCUS	AL578658/c					
DEFINITION	AL578658	Homo sapiens HELA CELLS	COT 25-NORMALIZED	Homo sapiens		
		cDNA clone CS0DK04Y114	3-PRIME,	mRNA sequence.		
ACCESSION	AL578658					
VERSION	AL578658.2	GI:31316853				
KEYWORDS	EST.					
SOURCE	Homo sapiens	(human)				
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 984)					

Li, W. B., Gruber, C., Jessee, J., and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On Feb 16, 2001 this sequence version replaced gi:12942945.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segref@genoscope.cns.fr, Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 5047.f For  
more information about this cluster, see  
<http://www.genoscope.cns.fr/>

```

FEATURES
source
1..984
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK004Y114"
cgi-bin/cluster.cgi?seq=CS0DK004BE07NP1&cluster=5047.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DK004BE07NP1.
Location/Qualifiers

```

```

/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="Not strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched double-strand cDNA was
digested with Not I and cloned into the Not I and EcoPv
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

```

	Query Match	88.2%;	Score 952.6;	DB 9;	Length 984;
	Best Local Similarity	98.4%;	Pred. No. 5.7e-236;		
	Matches 966;	Conservative 3;	Mismatches 12;	Indels 1;	Gaps 1;
QY	99	GGACGTAGCGCAAGCGCATCGCATCGCCATGGCCAAAGGAGTGTGGGGCTGATTGTGCCAAGTT	158		
Db	984	GGATGTAGCGCAAGCGCATCGCATGGCCAAAGGAGTGTGGGGCTGATTGTGCCAAGTT	925		
QY	159	CCAGAAAGGTGAGCTAGAAATTCAAGTTTAATCGGAAAGCCTTTGGAGAGGCCATACACCTC	218		
Db	924	CCAGAAAGGTGAGCTAGAAATTCAAGTTTAATCGGAAAGCCTTTGGAGAGGCCATACACCTC	865		





```
QY 541 TTCTTGCAGTGTACAGCGCATACCCGCTCCAGCCTGAGGAGCTCAACCTGCGGGTCATC 600
Db 529 TTCTTGCAGTGTACAGCGCATACCCGCTCCAGCCTGAGGAGCTCAA-CTGCGGGTCATC 471
QY 601 TCGGAATATFCAAGAGCTCTTCTCTGACATTCATCCATAGGGTATTCTGGGCATGAACAGCG 660
Db 470 TCGGAATATFCAAGAGCTCTTCTCTGACATTCATCCATAGGGTATTCTGGGCATGAACAGCG 411
QY 661 ATAGCGATATCTGTGGCCGAGTGGCTCTGGGGCCAGGTTGTGGAACTGACATAACT 720
Db 410 ATAGCGATATCTGTGGCCGAGTGGCTCTGGGGCCAGGTTGTGGAACTGACATAACT 351
QY 721 TTGGCAAGACCTGGAAGGGAGTGACCACTCGGCTCGCTCGAGCTCGAGCTCGAGAACTGGCC 780
Db 350 TTGGCAAGACCTGGAAGGGAGTGACCACTCGGCTCGCTCGAGCTCGAGAACTGGCC 291
QY 781 GAGCTGGTGGCTGAGTGGCTCTTGTGGAGCGTGGCTGGGCTCCCAACCAAGCAGCTG 840
Db 290 GAGCTGGTGGCTGAGTGGCTCTTGTGGAGCGTGGCTGGGCTCCCAACCAAGCAGCTG 231
QY 841 CTGCCCTGTGAGATGGCTGCAATGAGAGCTGGGCAAGTCTGTGGTGGCCCAAGTGAA 900
Db 230 CTGCCCTGTGAGATGGCTGCAATGAGAGCTGGGCAAGTCTGTGGTGGCCCAAGTGAA 171
QY 901 ATTCGGAGGACCACTTCTAAACAATGACATGCTCACCGTGAAGTGGGTGAGCCCAA 960
Db 170 ATTCGGAGGACCACTTCTAAACAATGACATGCTCACCGTGAAGTGGGTGAGCCCAA 111
QY 961 GCCTATCTCTGAGACATCTTTAATCTAGTGGGCAAGAGTCTCTGGTCACTGTGTGAA 1020
Db 110 GCCTATCTCTGAGACATCTTTAATCTAGTGGGCAAGAGTCTCTGGTCACTGTGTGAA 51
QY 1021 GAGGATGACACCATCATGGAAGAAATGGTAGATATCATGCGCAAAAAAT 1070
Db 50 GAGGATGACACCATCATGGAAGAAATGGTAGATATCATGCGCAAAAAAT 1
```

## RESULT 4

```
AL578579/c
LOCUS AL578579 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DK001YM22 3-PRIME, mRNA sequence.
```

```
ACCESSION AL578579
VERSION AL578579.1 GI:12942789
KEYWORDS EST.
SOURCE Homo sapiens (human)
```

## ORGANISM

```
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 990)
Li W.B., Gruber C., Jessee J. and Polayes D.
Full-length cDNA libraries and normalization
Unpublished (2001)
```

## REFERENCE

```
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5047.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DK001BG11NP1&cluster=5047.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DK001BG11NP1.
Location/Qualifiers
```

## FEATURES

source

```
1..990
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK001YM22"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
```

/clone.lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

```
Query Match 85.0%; Score 917.8; DB 9; Length 990;
Best Local Similarity 96.6%; Pred. No. 6.2e-227;
Matches 958; Conservative 8; Mismatches 23; Indels 3; Gaps 3;

QY 1 ATGCGCTGGAGCTGAGCTGTCCCGGGCGCTGGTGGGGGCGGCAACACCCGTCCTTC 60
Db 990 ATGCGCTGGGGCTGAGCTTTGTCCGGGCGCTGGTGGGGGCGGCAACACCCGTCCTTC 931
QY 61 ATCATTTGGCGAGATCGGCCAGAACACACAGGGCGACCTGGACCTAGCCAAAGTCATATTC 120
Db 930 ATAAATTGCGAGATCGGCCAGAACACACAGGGCGACCTGGATGTAGCCAAGCCATAATC 871
QY 121 CGCATGGCCAGGAGTGTGGGCTGATTGTGCAAGTTCCAGAAAGTCCAGAAAGTCCAGTAATTC 180
Db 870 CGCATGGCCAGGAGTGTGGGCTGATTGTGCAAGTTCCAGAAAGTCCAGTAATTC 811
QY 181 AAGTTTAAATCGGAAAGCCCTTGGAGAGGCCATACCTCGAAAGCATTCCTGGGGGAAGCG 240
Db 810 AAGTTTAAATCGGAAAGCCCTTGGAGAGGCCATACCTCGAAAGCATTCCTGGGGGAAGCG 751
QY 241 TACGGGGAGCAAAACGACATCTGAGATTACGCATGACCAAGTACAGGAGTCCAGAGG 300
Db 750 TACGGGGAGCAAAACGACATCTGAGATTACGCATGACCAAGTACAGGAGTCCAGAGG 691
QY 301 TACGGCGGAGGTTGGGATCTTCTTCACTGCTCGCATGGATGGATGAGATGCGAGTTGAA 360
Db 690 TACGGCGGAGGTTGGGATCTTCTTAACT-CCTTGGCATGGATGGAGTTGGA 632
QY 361 TTCTTGCATGAATGAATGTTCCATTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420
Db 631 TTCTTGCATGAATGAATGTTCCATTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 572
QY 421 CCTTATCTGGAAGACAGACCCAAAAGTGGCCCAATGGTGTATCTCCAGTGGGATGACG 480
Db 571 CCTTATCTGGAAGACAGACCCAAAAGTGGCCCAATGGTGTATCTCCAGTGGGATGACG 512
QY 481 TCAATGGACACCATGACCAAGCAAGTTTTCAGATCTGAGAGCCCTCAACCCCAACTCTGC 540
Db 511 TCAATGGACACCATGACCAAGCAAGTTTTCAGATCTGAGAGCCCTCAACCCCAACTCTGC 452
QY 541 TTCTTGCAGTGTACAGGCGCATACCCGCTCCAGCCTGAGGAGCTCAACCTGCGGGTCATC 600
Db 451 TTCTTGCAGTGTACAGGCGCATACCCGCTCCAGCCTGAGGAGCTCAAC-CTGCGGGTCATC 393
QY 601 TCGGAATATCAGAAGCTCTTCTCTGACATTTCCCATAGGGTATTCTGGGCATGAACAGGC 660
Db 392 TCGGNATNTCAGAAGCTCTTCTCTGACATTTCCCATAGGGTATTCTGGGCATGAGACAGGC 333
QY 661 ATAGCGATATCTGTGGCGCAGTGGCTCTGGGGGCGCAAGTGTGTGGAACTGACATACT 720
Db 332 ATAGCGATATCTGTGGCGGAGTGGCTCTGGGGGCGCAAGTGTGTGGAACTGACATACT 273
QY 721 TTGGAACAAGACCTGGAAGGGAGTGACCACTCGGCTCGCTGGAGCCCTGGAGAACTGGCC 780
Db 272 TTGGAACAAGACCTGGAAGGGAGTGACCACTCGGCTCGCTGGAGCCCTGGAGAACTGGCC 213
QY 781 GAGCTGGTGGCTGAGTGGCTCTTGTGAGAGCTGGCTGGGCTCCCAACCAAGCAGCTG 840
Db 212 GGGTGGTGGCTGAGTGGCTCTTGTGAGAGCTGGCTGGGCTCCCAACCAAGCAGCTG 153
QY 841 CTGCCCTGTGAGATGGCTTGAATGAGAAAGTGGGCAAGTCTGTGGTGGCCCAAGTGAAA 900
Db 152 CTGCCCTGTGAGATGGCTTGAATGAGAAAGTGGGCAAGTCTGTGGTGGCCCAAGTGAAA 93
QY 901 ATTCGGAGGACCACTTCTAAACAATGACATGCTCAACCGTGAAGTGGGTGAGC-CCAA 959
```

Db 92 ATTCGGGAAGCACCATTCTWACATGAACATGCTCACCGTGAAGGTGGGTGAGCGCCAA 33

Qy 960 AGCCTATCTCTCTGAAGACATCTTTAACTAG 991

Db 32 AGGCTATCTCTCTGAAGACATCTTTAACTAG 1

RESULT 5

AY414839

LOCUS

DEFINITION Mus musculus NANS gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION AY414839

VERSION AY414839.1 GI:39770798

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1080)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 1080)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES

source 1..1080

gene /organism="Mus musculus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10090"

<1..>1080

/gene="NANS"

/locus\_tag="HWM5357"

ORIGIN

Query Match 81.8%; Score 883.2; DB 29; Length 1080;

Best Local Similarity 88.6%; Pred. No. 6.3e-218;

Matches 957; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

Qy 1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGGGCGGCAACACCCGTGCTTC 60

Db 1 ATGCCGCTGGAACTGGAGCTGTGTCCCGGGCGCTGGGTGGGTGGAAAGCACCCGTGCTTC 60

Qy 61 ATCATTCGGAGATCGGCCAGAACACACAGGGCGACCTGGACGTAGCCAAAGCGATGATC 120

Db 61 ATCATTCGGAGATCGGCCAGAACACACAGGGAGACATAGATGTGGCCAAAGCGATGATC 120

Qy 121 CGCATGGCCAAAGGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAGAGTGAGTAGAATTC 180

Db 121 CGCATGGCCAAAGGAGTGTGGGGCGACTGCGTAAAGTTTCAAGAGCGAGTTGAGTTC 180

Qy 181 AAGTTTAACTCGAAAGCCTTGGAGAGCCATACACCTCGAAGCAATTCCTGGGGGAAGACG 240

Db 181 AAGTTTAACTCGAAAGCCTTGGAGAGCCATATATCTTCGAAGCAATTCATGGGGGAAGACG 240

Qy 241 TACGGGAGACCAACGACATCTGGAGTTGACCATGACCATGACAGTACAGGAGCTGCAGAGG 300

Db 241 TATGGGAGACCAAGCGGCATCTTGGAAATTCAGCCACGACAGTACAAAGAGCTGCAGAGC 300

Qy 301 TACGCCGAGGAGGTGGGATCTTCTTCACTCGCTCTGGCATGGATGAGATGCGAGTTGAA 360

Db 301 TATGGCGAGGAGATCGGCAATCTTCTTCACTGCTCTGGCATGGAGAGATGCACTTGAG 360

Qy 361 TTCTGTCATGAACCTGAATGTTCCATTTTCAAAGTTGGATCTGGAGACACTAATAATTTT 420

Db 361 TTCTGTCGCAAGAACTGAATGTTCCCTTTTCAAAGTTGGATCTGGGGACACTAACAATTTT 420

Qy 421 CTTATCTTGGAAAAGACAGCCAAAGGTGCCCAATGGTGATCTCCAGTGGGATGGAG 480

Db 421 CCCTACCTTGGAAAAGACAGCCAAAGGTGCTCTTATGTTGATCTCCAGCGGATGAG 480

Qy 481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTGC 540

Db 481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTGAAATCCCAACTTCTGC 540

Qy 541 TTCTTGAGTGTACAGCGCATACCCGCTCCAGCCTCGAGGACGTCAACCTCGGGGTCACTC 600

Db 541 TTCTTCCAATGCACAGCGCGTACCCACTACAGCCCGAGGATGCCAACCTTGGCGTCACTC 600

Qy 601 TCGGAATATCAGAACTCTTTCCTTGACATTCCTCATAGGGTATTTCTGGGCATGAACAGGC 660

Db 601 TCGGAATATCAGAACTCTTTCCTTGACATTCCTCATAGGGTATTTCTGGGCATGAACAGGC 660

Qy 661 ATAGCGATATCTGTGGCGCAGTGTCTCTGGGGCCCAAGGTGTTGGACAGTCAACATACT 720

Db 661 ATGCCATATCTGTGGCGCGCTGTCTGGGGCCCAAGGTGTTGGACAGTCAACATACT 720

Qy 721 TTGGACAAGACCTGGAAAGGGAGTGACCACTCGGCTCGCTGGAGCCTCGGAGAACTGGCC 780

Db 721 TTGGACAAGACCTGGAAAGGGAGTGACCACTCGGCTCGCTGGAGCCTCGGAGAACTGGCC 780

Qy 781 GAGCTGTGCGGTGAGTGTCTTGTGGAGGTGCTCTGGGGTCCCTGGGCTCCCAACCAAGCAGTG 840

Db 781 GAGCTGTGCGGTGAGTGTCTTGTGGAGGTGCTCTGGGGTCCCTGGGCTCCCAACCAAGCAGTG 840

Qy 841 CTGCCCTGTGAGATGGCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCCAAGTGAAG 900

Db 841 CTGCCCTGTGAGATGGCTGCAATGAGAAGCTGGGCAAGTCTGTGGTGGCCCAAGTGAAG 900

Qy 901 ATCCGGAAGGCACCACTTCTAACCAATGGACATGTCTACCGTGAAGGTGGGTGAGCCCAA 960

Db 901 ATCCGGAAGGCACCACTTCTAACCAATGGACATGTCTACCGTGAAGGTGGGTGAGCCCAA 960

Qy 961 GCCTATCTCTCTGAAGACATCTTTAATCTAGTGGCAAGAGGTCTCTGGTCACTGTTGAA 1020

Db 961 GCCTATCTCTCTGAAGACATCTTTAATCTAGTGGCAAGAGGTCTCTGGTCACTGTTGAA 1020

Qy 1021 GAGATGACACCATCATGGAAGAATTGTTAGATAATCATGCAAAAATAATCAAGTCTTAA 1080

Db 1021 GAGATGACACCATCATGGAAGAATTGTTAGATAATCATGCAAAAATAATCAAGTCTTAA 1080

RESULT 6

AK076290

LOCUS

DEFINITION Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched library, clone:463241B504 product:sialic acid synthase, full insert sequence.

ACCESSION AK076290

VERSION AK076290.1 GI:26345279

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1

AUTHORS Carninci,P. and Hayashizaki,Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2

AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,

Itch,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159

# REFERENCE

3  
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,  
Konno,H., Akiyama,J., Nishi,K., Kitsuana,T., Tashiro,H., Itoh,M.,  
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,  
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,  
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,  
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,  
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

# REFERENCE

4  
The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)  
5  
The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 1906)

# REFERENCE

Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,  
Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,  
Furuno,M., Hanagaki,F., Hara,A., Hayatsu,N., Hiramoto,K.,  
Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,  
Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,  
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,  
Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,  
Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,  
Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,  
Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,  
Tejima,Y., Toya,T., Yamamura,T., Yamana,I., Yasunishi,A.,  
Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.  
Direct Submission

# TITLE

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail:genome-res@sc.riken.go.jp,  
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,  
Fax:81-45-503-9216)

# COMMENT

cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL:http://genome.gsc.riken.go.jp/  
URL:http://fantom.gsc.riken.go.jp/

# FEATURES

source  
1. .1906  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:4632418E04"  
/db\_xref="MGI:1903268"  
/db\_xref="taxon:10090"  
/clone="4632418E04"  
/tissue\_type="skin"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="0 day neonate"  
32. .1111  
/note="unnamed protein product; putative"

# CDS

sialic acid synthase (MGI|2149820, GB|NM\_053179,  
evidence: BLASTN, 100%, match=1140)"  
/codon\_start=1  
/protein\_id="BAC36290.1"  
/db\_xref="GI:26345280"  
/translation="MPLELECPGRWVGKHPKCFIIAIGONHQHDIIDVAKRMRTAK  
ECGADCAKFKQSELEFKENRKALEPPYTSKHSWGKTYGEHKHLEHFSHDYKELOSVA  
QEMTGIFFTASGDEMAVEFLHNLNVPFFKVGSGDNNPYLEKTKAKRPRWISGMQ  
SMDTKQYQIVKPLNPFCLQCTSAYPLOPEDANLRVISEYQKLPDIPDIGSGHP  
TGIAISVAALGAKLVERHHTLDKTKWGSDBSLEFGELEAELVRSVRLVRLALGSP  
TKQLLPCSMACNEKLGKSVAKVKIPAGTTTLTLDMLTVKVGEKPKYPEDIPNLAKKK  
VLVTIEEDTWMESVESHSKKIKA"

# ORIGIN

Query Match	81.8%;	Score	883.2;	DB 11;	Length	1906;
Best Local Similarity	88.6%;	Pred. No.	8.5e-218;			
Matches	957;	Conservative	0;	Mismatches	123;	Indels
						0; Gaps
						0;

  

QY	1	ATGCGCTGGAGCTGGAGCTGTGTCCTCGGGCGCTGGTGGGGCGGCAACACCCGTCCTTC	60
DB	32	ATGCCCTGGAACTGGAGCTGTGTCCTCGGGCGCTGGTGGGGCGGCAACACCCGTCCTTC	91
QY	61	ATCATTTGCCGATCGGCCAGAACACACAGGGCGACCTGGAGCTAGCCAGCGCATGATC	120
DB	92	ATCATCGCGGATCGGCCAGAACACACAGGGAGACATAGATGTGCCAAGCGCATGATC	151
QY	121	CGCATGGCCAGGAGTGGGGCTGATTTGTGCCAAGTTCCAGAGAGTGGAGTAGATTC	180
DB	152	CGCACTGCCAAGGAGTGGGGCGGCTGCGCTAAGTTTCAGAGAGCGAGTTGGAGTTC	211
QY	181	AAGTTTAATCGGAAGCCCTTGGAGAGGCGATACACCTCGAAGCATTCCTGGGGGAAGACG	240
DB	212	AAAGTTTACCGGAGGCCCTGGAGAGACATATATCTTCAGAGCATTCATGGGGGAGACG	271
QY	241	TACGGGAGCAAAACGACATCTGGAGTTCAGCCATGACCATGACAGGAGCTGCAGAGG	300
DB	272	TATGGGAGCAACAGCGCATCTGGAATTCAGCCACGACCATGATCAAGGAGCTGCAGAGC	331
QY	301	TACGGGAGGAGTGGGATCTTCTTCACTGCTCTGSCATGGATGAGATGTCAGTTGNA	360
DB	332	TATGGGAGGAGTGGGATCTTCTTCACTGCTCTGSCATGGAGAGATGTCAGTTGAG	391
QY	361	TTCTGTCATGAATGAATGTTCCATTTTCAAAGTTGGATCTGGAGACACATAATATTTT	420
DB	392	TTTCTGACGAACTGAATGTTCCCTTTTCAAAGTTGGATCTGGGGACACATAACATTTT	451
QY	421	CCTTATCTGGAAGACAGCAAAAGAGTGGCCCAATGGTGCCTCAGTGGGATGCGAG	480
DB	452	CCCTACCTGGAAGACAGCAAAAGAGTGGTCTCTCTGCTGCTGAGCGGATGCGAG	511
QY	481	TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCACTTCCTG	540
DB	512	TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAATCCCACTTCCTG	571
QY	541	TTCTTGCAGTGTACCAAGCGCATACCCGCTCCAGCTGAGGAGCTCAACTCGCGGCTCATC	600
DB	572	TTCTTCCAAATGACCAAGCGCGTACCCACTACAGCCCGAGGATGCCAACTGCGCGTATC	631
QY	601	TCGGAATATCAGAGCTTTTCTGTCGATTCATAGGGTATTCCTGGGATGAAGACAGC	660
DB	632	TCGGAATATCAGAGCTTTTCTGTCGATTCATAGGGTATTCCTGGGATGAAGACAGC	691
QY	661	ATAGCGATATCTGTCGCGCAGTGGCTCTGGGGCCAGAGTGTGGACGTACATAACT	720
DB	692	ATGCGCATATCTGTGGCCCGCTGGCTCTGGGGCCAGAGTGTGGACGTACATAACT	751
QY	721	TTTGAAGAAGCCTGGAAGGGAGTGACCACTCGGCTCGCTGGAGCTGGAGAACTGGCC	780
DB	752	TTTGAAGAAGCCTGGAAGGGAGTGACCACTCGGCTCGCTGGAGCTGGAGAACTGGCA	811
QY	781	GAGCTGGTGGCTGAGTGGCTTTGTGAGAGCGTGGCTCGGCTCGGCTCGGCTCGGCTG	840
DB	812	GAGCTGGTGGCTGAGTGGCTTTGTGAGAGCGTGGCTCGGCTCGGCTCGGCTCGGCTG	871

QY 841 CTGCCCTGTGAGATGCTGCAATGAGAGCTGGGCAAGCTCTGCTGGTCCCAAGTGAA 900  
 Db 872 CTGCCCTGTGAGATGCTGCAATGAGAGCTGGGCAAGCTCTGCTGGTCCCAAGTGAA 931  
 QY 901 ATTCCGAGAGGCCACCATCTTAACAATGGACATGCTCACCGTGAAGTGGGTGAGGCCCAA 960  
 Db 932 ATCCAGCAGGACCAACCTGACCTGGACATGCTCACTGTGAAGTGGGGAGGCCCAA 991  
 QY 961 GCCTATCTCTGAAGACATCTTTAATCTAGTGGGCAAGAAGTCTGGTCACTGTGTAA 1020  
 Db 992 GGTATCTCTCTGAAGACATCTTTAATCTAGTGGGCAAGAAGTCTGGTCACTGTGTAA 1051  
 QY 1021 GAGGATGACACCATCATGGAAGAAATTTGGTGTAGATATCATGGCAAAATCAAGTCTTAA 1080  
 Db 1052 GAAGATGACACGCTCATGGAAGAAATCCGTGGAAAGTTCACAGCAAGAAATCAAGGCTTAA 1111

RESULT 7  
 AK088912  
 LOCUS  
 DEFINITION Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-length enriched library, clone:B430030P04 product:sialic acid synthase, full insert sequence.  
 AK088912  
 ACCESSION AK088912.1 GI:26104987  
 VERSION  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1  
 Carninci, P. and Hayashizaki, Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636  
 2  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20499374  
 PUBMED 11042159  
 3  
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, I., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE 20530913  
 PUBMED 11076861  
 4  
 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
 TITLE Functional annotation of a full-length mouse cDNA collection  
 JOURNAL Nature 409, 685-690 (2001)  
 5  
 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
 TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 JOURNAL Nature 420, 563-573 (2002)  
 6 (Bases 1 to 2109)  
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,

Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
 Direct Submission  
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]  
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.  
 Please visit our web site for further details.  
 URL: http://genome.gsc.riken.go.jp/  
 URL: http://fantom.gsc.riken.go.jp/  
 Location/Qualifiers  
 1..2109  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="NOD"  
 /db\_xref="FANTOM DB:B430030P04"  
 /db\_xref="MGI:2427803"  
 /db\_xref="taxon:10090"  
 /clone="E430030P04"  
 /cell\_type="thymic cells"  
 /tissue\_type="thymus"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="2 days neonate"  
 misc\_feature  
 41..1121  
 /note="putative sialic acid synthase (MGD|MGI:2149820, GB|NM\_053179, evidence: BLASTN, 100%, match=1140)"

## ORIGIN

Query Match 80.6%; Score 870.6; DB 11; Length 2109;  
 Best Local Similarity 88.4%; Pred. No. 1.7e-214;  
 Matches 956; Conservative 0; Mismatches 124; Indels 1; Gaps 1;  
 QY 1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGCGGCAACACCCGCTGCTTC 60  
 Db 41 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCGCTGGGTGGCGGCAACACCCGCTGCTTC 100  
 QY 61 ATCATTCGCGAGATCGGCCAGAACCCACGAGCGGACCTGGACGTAGCCAGCGCATGATC 120  
 Db 101 ATCATTCGCGAGATCGGCCAGAACCCACGAGCGGACATAGATGTGGCCAGCGCATGATC 160  
 QY 121 CGCATGCGCAAGGAGTGTGGGGCTGATGTGCGAAGTTCCGAAAGAGTGAGCTAGATTTC 180  
 Db 161 CGCACTCGCAAGGAGTGTGGGGCGGCTGCGCTAAGTTTCAGAGAGAGAGTTGGAGTTC 220  
 QY 181 AAGTTTAAATCGGAAGCCTTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGAAGACG 240  
 Db 221 AAGTTTAAATCGGAAGCCTTGGAGAGGCCATATATCTTCGAAGCAATTCATGGGGGAAGACG 280  
 QY 241 TACGGGAGGACAAACGACATCTGGAGTTTCAGGCTACGACATACAGGAGCTCAGAGG 300  
 Db 281 TATGGGAGGACAAACGAGCATCTGGATTTCAGCCACGACCATACAGGAGCTCAGAGG 340  
 QY 301 TACCCGAGGAGGTGGGATCTTCTTCACTGCGCTCTGGCATGGATGAGATGGCAGTTGAA 360

Db 341 TATGGCAGGAGATCGGATCTTTCTCACTGCTCTGGATGACGAGATGCGAGTTCAG 400  
 Qy 361 TTCTGTCATGAACGATGAATGTTCCATTTTAAAGTTGGATCTGGAGACACATAATATTTT 420  
 Db 401 TTCTGTCAGCACTGAATGTTCCCTTTTCAAAGTTGGATCTGGGACACATAACAATTTT 460  
 Qy 421 CTTATCTGGAAAGACAGCCAAAAGGTCGCCCAATGTTGGTATCTTCAGTGGGATGCGAG 480  
 Db 461 CCTACCTGGAAAGACAGCCAAAAGGTCGCCCAATGTTGGTATCTTCAGTGGGATGCGAG 520  
 Qy 481 TCAATGGACACCATCAAGCAAGCTTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCGCG 540  
 Db 521 TCAATGGACACCATCAAGCAAGCTTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCGCG 580  
 Qy 541 TTCTGTCAGTGTACAGCGCATACCCGCTCCAGCCTGAGGACGTCGAACCTGCGGGTCATC 600  
 Db 581 TTCTCCTCAATGACACGCGGTACCCACTACAGCCCGAGGATGCCAACCTGCGGTCATC 640  
 Qy 601 TCGGAATATCAGAACTCTTTCTGACATTTCCATAGGATTTCTGGGATGAACACAGGC 660  
 Db 641 TCGGAATACCAAGAACTCTTTTCCGACATTTCCCATCGGGTATTTCCGGGACAGAGCGGC 700  
 Qy 661 ATAGCGATATCTGTGCGCGAGTGTCTGTGGGGCCAAAGTGTGTGGAAGCTCACATAACT 720  
 Db 701 ATCGGCATATCTGTGCGCGCGTGTGTGTGGGGCCAAAGTGTGTGGAAGCTCACATAAG 760  
 Qy 721 TTGGACAAGACTGGAAGGGGAGTGAACCACTCGGCTCTGAGGACCTGGAGAACTGGCC 780  
 Db 761 TTGGACAAGACTGGAAGGGGAGTGAACCACTCAGCTCTGAGGCTGGGGAGCTGGCA 820  
 Qy 781 GAGCTGGTTCAGTGGTCTTTGTGAGCGTGCCTGGGCTCCCAACCAAGCAGCTG 840  
 Db 821 GAGCTGGTTCAGTGGTCTTTGTGAGCGTGCCTGGGCTCCCAACCAAGCAGCTG 880  
 Qy 841 CTGCGCTGTGAGATGCGCTGCAATGAGAGCTGGGCAAGTCTGTGTGGGCCAAAGTGAAA 900  
 Db 881 CTGCGCTGTGAGATGCGCTGCAATGAGAGCTGGGCAAGTCTGTGTGGTAGCAAGTGAAA 940  
 Qy 901 ATTCCGGAAGGACCACTTCTAAATATGACATGCTCAGCGTGAAGGTTGGTGGAGCCCAA 960  
 Db 941 ATCCGAGAGGACCACTGACCTGAGCATGCTCACTGTGAAGTGGGGAGGCCCAA 1000  
 Qy 961 GCCTATCTCTGAGACATCTTAACTAGTGGGC-AGAGAGTCTGCTGCTGCTGGA 1019  
 Db 1001 GGCTATCTCTGAGACATCTTAACTAGTGGGC-AGAGAGTCTGCTGCTGCTGGA 1060  
 Qy 1020 AGAGATGACACCATCATGGAAGAATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTA 1079  
 Db 1061 AGAGATGACACCGTCATGGAAGATCCGTGGAAGTCAACACAGAAATCAAGCTTA 1120  
 Qy 1080 A 1080  
 Db 1121 A 1121

RESULT 8  
 B1517350  
 LOCUS  
 DEFINITION B1517350 925 bp mRNA linear EST 29-AUG-2001  
 603041586F1 NIH\_MGC\_116 Homo sapiens cDNA clone IMAGE:5163020 5',  
 mRNA sequence.  
 ACCESSION B1517350  
 VERSION B1517350  
 KEYWORDS B1517350.1 GI:15342142  
 EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 925)  
 NIH-MGC http://mgi.mci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.

Email: cgapbs-re@mail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1AM11404 row: k column: 21  
 High quality sequence stop: 856.  
 Location/Qualifiers  
 1. 925  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5163020"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH MGC 116"  
 /note="Organ: pooled colon, kidney, stomach; Vector:  
 pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA  
 source anonymous pool of 3 colons, age 26 yo male, 49 yo  
 female, 71 yo male colon; 46 yo male kidney, and pool of 2  
 stomachs, 62 yo male and 70 yo female. Library is  
 oligo-dT primed and directionally cloned (EcoRV site is  
 destroyed upon cloning). Average insert size 1.4 kb,  
 insert size range 1-3 kb. Library is normalized and  
 enriched for full-length clones and was constructed by C.  
 Gruber (Invitrogen). Research Genetics tracking code  
 023. Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 79.3%; Score 856.6; DB 12; Length 925;  
 Best Local Similarity 98.8%; Pred. No. 4.7e-211;  
 Matches 884; Conservative 0; Mismatches 9; Indels 2; Gaps 2;  
 Qy 188 ATCGAAAGCCCTGGAG-AGGCCATACACCTCGAAGCATTCCTGGGGGAAGACGTACGGG 246  
 Db 1 ATCGAAAGCCCTGGAGAGGCCATACACCTCGAAGCATTCCTGGGGGAAGACGTACGGG 60  
 Qy 247 GAGCAAAACGACATCTGGAGTTAGCCATGACCAAGTACAGGAGCTGCAGAGGTACGCC 306  
 Db 61 GAGCAAAACGACATCTGGAGTTAGCCATGACCAAGTACAGGAGCTGCAGAGGTACGCC 120  
 Qy 307 GAGGAGGTGGGATCTTCTTCACTGCTGCGCATGATGATGCGAGTGAATTCCTG 366  
 Db 121 GAGGAGGTGGGATCTTCTTCACTGCTGCGCATGATGATGCGAGTGAATTCCTG 180  
 Qy 367 CATGAACCTGAATGTTCCATTTTCAAAGTTGGATCTGGAGACACATAATAATTTCTTAT 426  
 Db 181 CATGAACCTGAATGTTCCATTTTCAAAGTTGGATCTGGAGACACATAATAATTTCTTAT 240  
 Qy 427 CTGGAAAAGACAGCCAAAAGGTCGCCCAATGTTGATCTCCAGTGGGATGCAAGTCAATG 486  
 Db 241 CTGGAAAAGACAGCCAAAAGGTCGCCCAATGTTGATCTCCAGTGGGATGCAAGTCAATG 300  
 Qy 487 GACACATGACAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCGCTCTTG 546  
 Db 301 GACACATGACAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCGCTCTTG 360  
 Qy 547 CAGTGTACAGCGCATACCCGCTCAGGCTGAGGACGTCAACCTCGGGTCAATCTCGGAA 606  
 Db 361 CAGTGTACAGCGCATACCCGCTCAGGCTGAGGACGTCAACCTCGGGTCAATCTCGGAA 420  
 Qy 607 TATCAGAGCTCTTCTTGACATTTCCCATAGGTTATCTGGGCATGAACAGGATAGCG 666  
 Db 421 TATCAGAGCTCTTCTTGACATTTCCCATAGGTTATCTGGGCATGAACAGGATAGCG 480  
 Qy 667 ATATCTGGGCGCAGTGGCTCTGGGGCCCAAGGTTGGACGTCAACATACTTTGGAC 726  
 Db 481 ATATCTGGGCGCAGTGGCTCTGGGGCCCAAGGTTGGACGTCAACATACTTTGGAC 540  
 Qy 727 AAGACCTGGAAGGGAGTGACCACTCGGCTCGCTGGAGCTGGAGAACTGGCCGAGCTG 786





TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT On Feb 15, 2001 this sequence version replaced gi:12898807.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 5047.f For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DA001BG1QPl&cluster=5047.f. Contact :  
Feng Liang Email: fliang@lifetech.com URL: Corporation 1600  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DK001BG1QPl.  
Location/Qualifiers  
1..1005  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DK001YM22"  
/cell\_type="HELA CELLS COT 25-NORMALIZED"  
/clone\_lib="HELA"  
/clone\_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES  
source  
1..1005  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DK001YM22"  
/cell\_type="HELA CELLS COT 25-NORMALIZED"  
/clone\_lib="HELA"  
/clone\_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
Query Match 78.1%; Score 843; DB 9; Length 1005;  
Best Local Similarity 95.5%; Pred. No. 1.7e-207; Indels 0; Gaps 0;  
Matches 846; Conservative 6; Mismatches 34; Indels 0; Gaps 0;  
QY 1 ATGCCGCTGGAGCTGTGTCTCCGGCGCTGGTGGCGGCGCAACACCCGTCCTC 60  
DB 120 ATGCCGCTGGAGCTGTGTCTCCGGCGCTGGTGGCGGCGCAACACCCGTCCTC 179  
QY 61 ATCATTCGCGAGATCGCCAGAACCAACAGGCGACCTGGAGCTAGCCAGGCGATGATC 120  
DB 180 ATCATTCGCGAGATCGCCAGAACCAACAGGCGACCTGGAGCTAGCCAGGCGATGATC 239  
QY 121 CCGATGCGCAAGAGTGTGGGCTGATTTGTCGCAAGTTCACAGAGTGTAGCTAGATTC 180  
DB 240 CCGATGCGCAAGAGTGTGGGCTGATTTGTCGCAAGTTCACAGAGTGTAGCTAGATTC 299  
QY 181 AAGTTTAATCGGAAGCCCTGGAGAGGCCATACACTCGAAGCATTCCTGGGGGAGAGC 240  
DB 300 AANTNTAATCCGAANCCCTTGAGANGCCATACACTCGAANCAATTCCTGGGGGAGAGC 359  
QY 241 TACGGGAGCACAAACGACATCTGGAGTTCAGCCATGACAGGAGCTGCAGAGG 300  
DB 360 TACGGGAGCACAAACGACATCTGGAGTTCAGCCATGACAGGAGCTGCAGAGG 419  
QY 301 TACGGGAGGAGTTCGGATCTTCTTCACTGCTCTGGCATGGATGGAGTGGAGTTGAA 360  
DB 420 TACCCGAGAGGTTNGATCTTCTTCACTGCTCTGGCATGGATGGAGTGGAGTTCGAA 479  
QY 361 TTCTCTGCATGAATGATGTTCCATTTTCAAAGTTGGATCTGGAGACACTAATATTTT 420  
DB 480 TTCTCTGCATGAATGATGTTCCATTTTCAAAGTTGGATCTGGAGACACTAATATTTT 539  
QY 421 CCTTATCTGGAAGACGCCAARAGTCCGCCAATGCTGATCTCCAGTGGGATGACG 480  
DB 540 CTTTATCTGGAAGACGCCAARAGTCCGCCAATGCTGATCTCCAGTGGGATGACG 599  
QY 481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTGC 540  
DB 600 TTAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTGC 659  
QY 541 TTCTTGCAGTGTACCAAGCGCATACCGCTCCAGCTGAGGACGTCAACCTCGGGGTATC 600  
DB 660 TTCTTGCAGTGTACCAAGCGCATACCGCTCCAGCTGAGGACGTCAACCTCGGGGTATC 719

QY 601 TCGGAATATCAGAAGCTCTTTCTTGACATTCCTAGGATTCCTGGGCATGAACAGGC 660  
DB 720 TCGGAATATCAGAAGCTCTTTCTTGACATTCCTAGGATTCCTGGGCATGAACAGGC 779  
QY 661 ATAGCGATATCTGTGGCGCAGTGGCTCTGGGGGCAAGAGTGTGGAAACGTACATAACT 720  
DB 780 ATAGCGATATCTGTGGCGCAGTGGCTCTGGGGGCAAGAGTGTGGAAACGTACATAACT 839  
QY 721 TTGGAACAAGACTGGAAAGGAGTGAACCACTCGGGCTCTGGAGCCCTGGAGAACTGGCC 780  
DB 840 TTGGAACAAGACTGGAAAGGAGTGAACCACTCGGGCTCTGGAGCCCTGGAGAACTGGCC 899  
QY 781 GAGCTGTGGCGTCACTGCTCTGTGGAGCGTGCCTGGGCTCCCAACCAAGCAGCTG 840  
DB 900 GAGCTGTGGCGTCACTGCTCTGTGGAGCGTGCCTGGGCTCCCAACCAAGCAGCTG 959  
QY 841 CTGCCCTCTGAGATGGCTTCAATGAGAAGCTGGGCAAGTCTGTGG 886  
DB 960 CTGCCCTCTGAGAGGCTTCAATGAGAAGCTGGGCAAGTCTGTGG 1005  
RESULT 11  
AL515470 1201 bp mRNA linear EST 09-MAY-2003  
LOCUS AL515470 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone  
DEFINITION CS0DA002YI24 5-PRIME, mRNA sequence.  
ACCESSION AL515470  
VERSION AL515470.2 GI:30489137  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1201)  
AUTHORS Li.W.B., Gruber,C., Jessee,J. and Polayes,D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT On Feb 13, 2001 this sequence version replaced gi:12778963.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 5047.f For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DA002BE12QPl&cluster=5047.f. Contact :  
Feng Liang Email: fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DA002BE12QPl.  
Location/Qualifiers  
1..1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DA002YI24"  
/tissue\_type="NEUROBLASTOMA"  
/clone\_lib="Homo sapiens NEUROBLASTOMA"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo (dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."

FEATURES  
source  
1..1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DA002YI24"  
/tissue\_type="NEUROBLASTOMA"  
/clone\_lib="Homo sapiens NEUROBLASTOMA"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo (dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."

ORIGIN  
Query Match 77.8%; Score 840; DB 9; Length 1201;  
Best Local Similarity 98.7%; Pred. No. 1.1e-206;  
Matches 859; Conservative 7; Mismatches 2; Indels 2; Gaps 2;  
QY 1 ATGCCGCTGGAGCTGGAGCTGTGTCTCCGGCGCTGGTGGCGGCGCAACACCCGTCCTC 60  
DB 103 ATGCCGCTGGAGCTGGAGCTGTGTCTCCGGCGCTGGTGGCGGCGCAACACCCGTCCTC 162

QY 61 ATCATTTGCCGAGATCGGCCAGAACACACAGGCGGACCTGGAGCTGACCCAGGCGATGATC 120  
 Db |||||  
 QY 121 CGCATGGCCAGGAGTGTGGGCTGATTTGTGCCAAGTTCCAGAAAGAGTGTAGTGAATTC 180  
 Db |||||  
 QY 223 CGCATGGCCAGGAGTGTGGGCTGATTTGTGCCAAGTTCCAGAAAGTGTAGTGAATTC 282  
 Db |||||  
 QY 181 AAGTTTAAATCGAAAGCCTTGGAGAGGCCATACACTCGAAGCATTTCTTGGGGGAAGAGC 240  
 Db |||||  
 QY 283 AAGTTTAAATCGAAAGCCTTGGAGAGGCCATACACTCGAAGCATTTCTTGGGGGAAGAGC 342  
 Db |||||  
 QY 241 TACGGGGAGCAGACAGATCTCGAGTTCAGCCATGACCATGACAGGAGTGTAGAGG 300  
 Db |||||  
 QY 343 TACGGGGAGCAGACAGATCTCGAGTTCAGCCATGACCATGACAGGAGTGTAGAGG 402  
 Db |||||  
 QY 301 TACGCCGAGGAGTGTGGATCTTTCTTCACTGCTCTGGCATGGATGAGATGGCAGTTGAA 360  
 Db |||||  
 QY 403 TACGCCGAGGAGTGTGGATCTTTCTTCACTGCTCTGGCATGGATGAGATGGCAGTTGAA 462  
 Db |||||  
 QY 361 TTCTGTGATGAATGAATGTTTCCATTTTCAAGTTGGATCTGGAGACACTAATTAATTTT 420  
 Db |||||  
 QY 463 TTCTGTGATGAATGAATGTTTCCATTTTCAAGTTGGATCTGGAGACACTAATTAATTTT 522  
 Db |||||  
 QY 421 CCTTATCTGGAAAGACAGACGACCAAAAGGTGCGCCAAATGGTGCATCTCCAGTGGGATGAG 480  
 Db |||||  
 QY 523 CCTTATCTGGAAAGACAGACGACCAAAAGGTGCGCCAAATGGTGCATCTCCAGTGGGATGAG 582  
 Db |||||  
 QY 481 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTGC 540  
 Db |||||  
 QY 583 TCAATGGACACCATGAAGCAAGTTTATCAGATCGTGAAGCCCTCAACCCCAACTTCTGC 642  
 Db |||||  
 QY 541 TTCTGTGATGAATGAAGTATACCCGTCCAGCTGAGGAGCGTCAACCTCGGGGTATC 600  
 Db |||||  
 QY 643 TTCTGTGATGAATGAAGTATACCCGTCCAGCTGAGGAGCGTCAACCTCGGGGTATC 702  
 Db |||||  
 QY 601 TCGGAATATCAGAGCTCTTCTTCCAGATTCCTTCCAGTTCCTGAGGAGTATCTGGGCATGAACAGGC 660  
 Db |||||  
 QY 703 TCGGAATATCAGAGCTCTTCTTCCAGATTCCTTCCAGTTCCTGAGGAGTATCTGGGCATGAACAGGC 762  
 Db |||||  
 QY 661 ATAGCGATATCTGTGGCGCAGTGTCTCTGGGGCCAAAGGTGTTGGAACGTCAACATTA 720  
 Db |||||  
 QY 763 ATAGCGATATCTGTGGCGCAGTGTCTCTGGGGCCAAAGGTGTTGGAACGTCAACATTA 822  
 Db |||||  
 QY 721 TTGACAAAGCTTGAAGGGAGTGAACCACTCGGCTCGCTGGAGCTTGGAGAACTGGCC 780  
 Db |||||  
 QY 823 TTGACAAAGA-CTGGAAGGGAGTGAACCACTCGGCTCGCTGGAGCTTGGAGAACTGGCC 881  
 Db |||||  
 QY 781 GAGCTGTGGCTCAGTGTGCTGTGTGGAGGCTGCTGGGCTCCCGAACCAAGCAGCTG 840  
 Db |||||  
 QY 882 GAGCTGTGGCTCAGTGTGCTGTGTGGAGGCTGCTGGGCTCCCGAACCAAGCAGCTG 940  
 Db |||||  
 QY 841 CTGCCCCTGTGAGATGGCTGCAATGAGAAG 870  
 Db |||||  
 QY 942 CTGCCCCTGTGAGAGG-CTGCAATGAGAG 970  
 Db |||||

RESULT 12  
 AL526241/c  
 LOCUS 989 bp mRNA linear EST 23-MAY-2003  
 DEFINITION AL526241 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
 cDNA clone CS0DC016YM22 3-PRIME, mRNA sequence.  
 ACCESSION AL526241  
 VERSION AL526241.2 GI:31064102  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1. (bases 1 to 989)  
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 Full-length cDNA libraries and normalization

JOURNAL COMMENT

Unpublished (2001)  
 On Feb 13, 2001 this sequence version replaced gi:12789734.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seq@genoscope.cns.fr, Web: www.genoscope.cns.fr  
 Library was constructed by Life technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 5047.f For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0DC016BG1NP16cluster=5047.f. Contact:  
 Feng Jiang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID: CS0DC016BG1NP1.  
 Location/Qualifiers  
 1. 989

FEATURES source

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DC016YM22"  
 /tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo (dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 76.9%; Score 830; DB 9; Length 989;  
 Best Local Similarity 93.0%; Pred. No. 3.9e-204;  
 Matches 872; Conservative 34; Mismatches 26; Indels 6; Gaps 4;  
 QY 126 GGCCAAGAGTGTGGGGCTGATTGTGCCAAGTTCCAGAAAGAGTGTAGTGAATTCAGATT 185  
 Db |||||  
 QY 940 KGCCAAGGAGTGTGGGCTGATGTG---CTAAGTCCAGAGAGTGTAGTGAATTCAGATT 884  
 Db |||||  
 QY 186 TAATCGGAAGCCTTGGAGAGGCCATACACTCGAAGCATTTCTTGGGGGAAGAGCTACGG 245  
 Db |||||  
 QY 883 TAATCGGAAGCCTTGGAGAGGCCATACACTCGAAGCATTTCTTGGGGGAAGAGCTACGG 824  
 Db |||||  
 QY 246 GGACCAACAAGCATCTGGAGTTCAGCCATGACAGGTACAGGAGTGTGAGAGTACG 305  
 Db |||||  
 QY 823 GGACCAACAAGCATCTGGAGTTCAGCCATGACAGGTACAGGAGTGTGAGAGTACG 764  
 Db |||||  
 QY 306 CGAGGAGTGTGGGATCTTTCTTCACTGCTCTGGCATGGATGAGATGGAGTTCCT 365  
 Db |||||  
 QY 763 CGAGGAGTGTGGGATCTTTCTTCACTGCTCTGGCATGGATGAGATGGAGTTCCT 704  
 Db |||||  
 QY 366 GCATGAAGTGAATGTTCCATTTTCAAAGTTGGATCTGGAGACACTAATAATTTTCTTT 425  
 Db |||||  
 QY 703 GCATGAAGTGAATGTTCCATTTTCAAAGTTGGATCTGGAGACACTAATAATTTTCTTT 644  
 Db |||||  
 QY 426 TCTGGAAGAGACAGCAAAAGGTGCGCCAAATGCTGATCTCCAGTGGGATGCAATCAAT 485  
 Db |||||  
 QY 643 TTTGGAAGAGACAGCAAAAGGTGCGCCAAATGCTGATCTCCAGTGGGATGCAATCAAT 584  
 Db |||||  
 QY 486 GGACCATGAAGCAAGTTTATCAGATCTGGAAGCCCTCAACCCCAACTTCTGCTCTT 545  
 Db |||||  
 QY 583 GGACCATGAAGCAAGTTTATCAGATCTGGAAGCCCTCAACCCCAACTTCTGCTCTT 524  
 Db |||||  
 QY 546 GCATGTACAGAGCCATACCCGCTCCAGCTGAGGAGTCAACTCTGGGGTCAATCTCGGA 605  
 Db |||||  
 QY 523 GCATGTACAGAGCCATACCCGCTCCAGCTGAGGAGTCAACTCTGGGGTCAATCTCGGA 465  
 Db |||||  
 QY 606 ATATCAGAAGCTCTTTCTTGCATCTCCATAGGATTTCTGGGCATGAACAGCATAGC 665  
 Db |||||  
 QY 464 ATATCAGAAGCTCTTTCTTGCATCTCCATAGGATTTCTGGGCATGAACAGCATAGC 405  
 Db |||||  
 QY 666 GATATCTGTGGCGCAGTGTGCTCTGGGGGCAAGGTGTTGGAACGTCAATAACTTTTGA 725  
 Db |||||  
 QY 404 GATATCTGTGGCGCAGTGTGCTCTGGGGGCAAGGTGTTGGAACGTCAATAACTTTTGA 345  
 Db |||||  
 QY 726 CAAGACCTGGAAGGGGAGTG-ACCACTGCGCCTCGCTGGAGCCTGGAGAACTGGCCGAGC 784  
 Db |||||



JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs@mail.nih.gov  
Tissue Procurement: ATCC/DCTD/DTP  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLNL3530 row: e column: 10  
High quality sequence stop: 711.  
Location/Qualifiers  
1. .872  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6167793"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 72"  
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dn.  
Average insert size 2 kb. Library constructed by Life  
Technologies."

## ORIGIN

Query Match 74.4%; Score 803.2; DB 13; Length 872;  
Best Local Similarity 99.0%; Pred. No. 3.3e-197;  
Matches 908; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCTGGTGGCGGGCAACACCCGCTGCTC 60  
DB 57 ATGCCGCTGGAGCTGGAGCTGTGTCCCGGGCTGGTGGCGGGCAACACCCGCTGCTC 116

QY 61 ATCATTCGGAGATCGGCAGAACACACAGCGGACCTGGAGTCCAGCCAGCGATGATC 120  
DB 117 ATCATTCGGAGATCGGCAGAACACACAGCGGACCTGGAGTCCAGCCAGCGATGATC 176

QY 121 CGCATGGCAAGAGGTGGGGCTGATGTGCGCAAGTTCAGAGAGTGGAGTGAATTC 180  
DB 177 CGCATGGCAAGAGGTGGGGCTGATGTGCGCAAGTTCAGAGAGTGGAGTGAATTC 236

QY 181 AAGTTTAAATCGAAAGCTTGGAGAGCGCATACCTCGAAGCATTCCTGGGGGAGACG 240  
DB 237 AAGTTTAAATCGAAAGCTTGGAGAGCGCATACCTCGAAGCATTCCTGGGGGAGACG 296

QY 241 TACGGGAGCAACACGATCTGGAGTTCAGCCATGACAGTACAGGAGCTGCAGAGG 300  
DB 297 TACGGGAGCAACACGATCTGGAGTTCAGCCATGACAGTACAGGAGCTGCAGAGG 356

QY 301 TACGCCGAGGAGTGGGATCTTCTCACTGCTCTGCGATGGATGAGATGGAGTTGAA 360  
DB 357 TACGCCGAGGAGTGGGATCTTCTCACTGCTCTGCGATGGATGAGATGGAGTTGAA 416

QY 361 TTCTGATGAACTGAATGTTCCATTTTCAAAGTTGGATCTGGAGACCTAATAATTTT 420  
DB 417 TTCTGATGAACTGAATGTTCCATTTTCAAAGTTGGATCTGGAGACCTAATAATTTT 476

QY 421 CCTTATCTGGAAGACAGCCAAAGGTGCCCAATGGTGTCTCCAGTGGGATCGAG 480  
DB 477 CCTTATCTGGAAGACAGCCAAAGGTGCCCAATGGTGTCTCCAGTGGGATCGAG 536

QY 481 TCAATGGACACCATGAGCAAGTTTATCAGATCTGAGCCCTCAACCCCACTTCTGC 540  
DB 537 TCAATGGACACCATGAGCAAGTTTATCAGATCTGAGCCCTCAACCCCACTTCTGC 596

QY 541 TTCTTGAGTGTACAGCGCATACCCGCTCCAGCTCGAGGACGTCAACCTCGGGGTATC 600  
DB 597 TTCTTGAGTGTACAGCGCATACCCGCTCCAGCTCGAGGACGTCAACCTCGGGGTATC 656

QY 601 TCGAATATCGAAGCTCTTTCCTGAATTCCTCCATAGGTTATCTGGGCGATGAACAGGC 660

DB 657 TCGGAATATCAGAGCTCTTCTTGACATTCCTAGGGTAATTCCTGGGCATGAACAGGC 716  
QY 661 ATAGCGATATCTGTGGCCGAGTGGCTTGGGGGCCAAGGTGTGGAACTGCACATAACT 720  
DB 717 ATAGCGATATCTGTGGCCGAGTGGCTTGGGGGCCAAGGTGTGGAACTGCACATAACT 776  
QY 721 TTGGCAAGACCTGGAGGGAGTACCACTGGCTCTCGCTGGAGCTGGAGAACTGGCC 780  
DB 777 TTGGCAAGACCTGGAGGGAGTACCACTGGCTCTCGCTGGAGCTGGAGAACTGGCC 836  
QY 781 GAGCTGGTGGGTGAGTGGCTCTTGTGGAGCTGGCC 816  
DB 837 GAGCTGGTGGGTGAGTGGCTCTTGTGGAGCTGGCC 872

RESULT 15  
LOCUS

DEFINITION BQ064521 1031 bp mRNA linear EST 02-APR-2002  
AGENCOURT\_6853542 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:5926258  
5', mRNA sequence.

## ACCESSION

VERSION BQ064521  
BQ064521.1 GI:19893231

## KEYWORDS

EST.

## SOURCE

Homo sapiens (human)

## ORGANISM

Homo sapiens

## REFERENCE

AUTHORS NIH-MGC

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs@mail.nih.gov  
Tissue Procurement: Lou Staudt

cDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCM2099 row: e column: 11

High quality sequence start: 22

High quality sequence stop: 706.

## FEATURES

source

1. .1031

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5926258"

/tissue\_type="lymphoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_99"

/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:

EcoRI; cDNA made by oligo-dT priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GGCAGCAG(G). Size-selected >500bp for average insert size

1.8kb. Library constructed by Ling Hong in the laboratory

of Gerald M. Rubin (University of California, Berkeley)

using ZAP-cDNA synthesis kit (Stratagene) and Superscript

II RT (Life Technologies). Note: this is a NIH\_MGC

Library."

## ORIGIN

Query Match 73.7%; Score 796.4; DB 13; Length 1031;  
Best Local Similarity 99.1%; Pred. No. 2.1e-195;  
Matches 811; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 264 GGAGTTGAGCATGACCAAGTACAGGGAGTGCAGAGGTACCCGAGGAGGTGGGATCTT 323  
DB 23 GGAGTTGAGCATGACCAAGTACAGGGAGTGCAGAGGTACCCGAGGAGGTGGGATCTT 82  
QY 324 CTTCACTGCTCTGGCATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGAT 383

Db 83 CTTCACTGCTCTGGCATGGATGAGATGGCAGTTGAATTCCTGCATGAATGTTCC 142  
Qy 384 ATTTTCAAAGTTGGATCTGGAGACATAATAATTTTCCTTATCTGGAAAAAGACAGCCAA 443  
Db 143 ATTTTCAAAGTTGGATCTGGAGACATAATAATTTTCCTTATCTGGAAAAAGACAGCCAA 202  
Qy 444 AAAAGTTCGCCAATAGTGTATCTCCAGTGGGATGAGTCAATGGACACCATGAAGCAAGT 503  
Db 203 AAAAGTTCGCCAATAGTGTATCTCCAGTGGGATGAGTCAATGGACACCATGAAGCAAGT 262  
Qy 504 TTATCAGATCTGAAGCCCTCAACCCCAAATTCTGTCTTTTGGAGTGTACAGCGCATA 563  
Db 263 TTATCAGATCTGAAGCCCTCAACCCCAAATTCTGTCTTTTGGAGTGTACAGCGCATA 322  
Qy 564 CCGCTCCAGCTGAGGACGTCAACCTCGGGTCAATCTCGGAATATCAAGAGTCTTTTCC 623  
Db 323 CCGCTCCAGCTGAGGACGTCAACCTCGGGTCAATCTCGGAATATCAAGAGTCTTTTCC 382  
Qy 624 TGACATTCCTCATAGGTATCTGGGCATGAACAGGCATAGCGATATCTGTGGCCGCGAGT 683  
Db 383 TGACATTCCTCATAGGTATCTGGGCATGAACAGGCATAGCGATATCTGTGGCCGCGAGT 442  
Qy 684 GGCTCTGGGGCCAAAGTGTGGAAAGTCAACATACTTTGGACAAGACCTGGAAGGGAG 743  
Db 443 GGCTCTGGGGCCAAAGTGTGGAAAGTCAACATACTTTGGACAAGACCTGGAAGGGAG 502  
Qy 744 TGACCATCGGCTCGCTGGAGCTTGAGAACTGGCCGAGCTGGTGGGTCAAGTGGTCT 803  
Db 503 TGACCATCGGCTCGCTGGAGCTTGAGAACTGGCCGAGCTGGTGGGTCAAGTGGTCT 562  
Qy 804 TGTGAGAGTGCCTGGGCTCCCAACCAAGCAGCTGTGCCCTGTGAGATGGCTTGCAA 863  
Db 563 TGTGAGAGTGCCTGGGCTCCCAACCAAGCAGCTGTGCCCTGTGAGATGGCTTGCAA 622  
Qy 864 TGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAATAATTCGGAAGGCAACATTCTAAC 923  
Db 623 TGAGAAGCTGGGCAAGTCTGTGGTGGCCAAAGTGAATAATTCGGAAGGCAACATTCTAAC 682  
Qy 924 AATGACATGCTCACCGTGAAGGTGGTGGCCCAAGCCTATCTCTGGAAGACATCTT 983  
Db 683 AATGACATGCTCACCGTGAAGGTGGTGGCCCAAGCCTATCTCTGGAAGACATCTT 742  
Qy 984 TAATCTAGTGGGCAAGAGTCTGTGTCACTGTGAAGAGGATGACACCATCATGG-AAAG 1042  
Db 743 TAATCTAGTGGGCAAGAGTCTGTGTCACTGTGAAGAGGATGACACCATCATGGAAAG 802  
Qy 1043 AATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 1080  
Db 803 AATTGGTAGATAATCATGGCAAAAAAATCAAGTCTTAA 840

Search completed: September 16, 2004, 05:19:48  
Job time : 4423 secs